The use of machine learning to develop a credit card fraud detection model for financial institutions.

## Author: Marta Elena Serrano Delgado

A Thesis Submitted in Partial Fulfilment

of the requirements for the

Degree of

Master of Science in Data Analytics



September 2023

Supervisor: David McQuaid

**CCT College Dublin**

**Assessment Cover Page**

*To be provided separately as a Word doc for students to include with every submission.*

|  |  |
| --- | --- |
| **Module Title:** | Capstone Project |
| **Assessment Title:** | The use of machine learning to develop a credit card fraud detection model for financial institutions. |
| **Supervisor Name:** | David McQuaid |
| **Student Full Name:** | Marta Elena Serrano Delgado |
| **Student Number:** | 2022300 |
| **Assessment Due Date:** | 24th of September 2023 |
| **Date of Submission:** | 24th of September 2023 |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

# Abstract

Internet banking and online shopping have grown significantly in today's world; therefore, the risk of fraudulent actions is studied: because of the number of fraud cases in daily transactions. Credit card fraud is becoming a higher problem for banks in recent years, resulting in financial losses on a global scale. When an unauthorised individual uses another person's credit card details to make transactions, this is referred to as credit card fraud, however, there are many cases when the credit card holder doesn’t recognise in an early stage the transaction until receives a notification from the financial institution and probably are more than one operation already. Credit card fraud is a significant and growing issue for institutions and people worldwide, this is the reason why the following project was considered. This thesis applies different techniques and models for machine learning in order to detect fraudulent transactions in an initial act. The datasets applied were obtained from different platforms: the first dataset is from Datacamp, which contains a total of 339607 rows from which 337825 are legit and 1782 are fraudulent transactions with a total of 15 features per operation, the second dataset is from Deloitte with a total of 6362620 rows, from which 6354407 are legit and 8213 are fraudulent transactions with a total of 10 features per transaction.

In the first analysis, the supervised models performed reasonably. The XGBoost Classifier Model performed the best with a perfect accuracy of 100%, however, it’s important to mention that this is a model trained with label data. Under-sampling was functional as part of the approaches in the second analysis, and the model with the best performance was the Random Forest Classifier, which did an excellent performance with an exceptional accuracy of 99%, despite only 10515 entries being trained, this because entire dataset contains a massive quantity of transactions, making the model difficult to perform in terms of time and computation, overall, the results obtained were better apply.

# Acknowledgments and Dedication

I am grateful to my supervisor, David McQuaid, who was tremendously helpful throughout the master's program and especially throughout the thesis process.

I am also grateful to all of the lecturers who provided me with the opportunity to learn, experience, and increase my knowledge during the several semesters at CCT College Dublin.

I dedicate this study to Gabriel, my wonderful love who has always been at my side, bringing forth the best in me. My family is my unwavering support, even if I am far away. All my family and friends play a significant role in shaping who I am and who I aspire to be.

LIST OF CONTENT

[CHAPTER 1 7](#_Toc146485323)

[1.1 Introduction 7](#_Toc146485324)

[1.2 Research Question 8](#_Toc146485325)

[1.3 Relevance 8](#_Toc146485326)

[1.4 Contribution 8](#_Toc146485327)

[1.5 Objectives 9](#_Toc146485328)

[CHAPTER 2 10](#_Toc146485329)

[2.1 Literature Review 10](#_Toc146485330)

[2.1.1 Credit card fraud application and findings. 10](#_Toc146485331)

[2.1.2 Types of electronic frauds. 13](#_Toc146485332)

[2.1.3 Different techniques used by credit card fraudsters. 14](#_Toc146485333)

[2.1.4 Difficulties of credit card fraud detection. 15](#_Toc146485334)

[Credit card fraud detection techniques. 15](#_Toc146485335)

[Dealing with Fraud Detection Approaches. 16](#_Toc146485336)

[2.1.5 Supervised Machine Learning. 16](#_Toc146485337)

[2.1.6 Clustering Methods. 17](#_Toc146485338)

[Imbalanced Dataset 17](#_Toc146485339)

[Resampling approach 17](#_Toc146485340)

[SMOTE stands for Synthetic Minority Over-sampling Technique. 17](#_Toc146485341)

[SMOTE and Tomek Link removal combined. 18](#_Toc146485342)

[2.1.7 HMM (Hidden Markov Model). 18](#_Toc146485343)

[2.1.8 SVM stands for Support Vector Machine. 19](#_Toc146485344)

[The Decision Tree. 19](#_Toc146485345)

[Algorithm Genetic. 19](#_Toc146485346)

[The Artificial Neural Network. 19](#_Toc146485347)

[Recurrent and LSTM Neural Networks. 20](#_Toc146485348)

[2.2 Validity 21](#_Toc146485349)

[2.3 Sampling Strategy 21](#_Toc146485350)

[2.4 Primary Research and Ethics. 22](#_Toc146485351)

[CHAPTER 3 22](#_Toc146485352)

[3.1 Methodology 22](#_Toc146485353)

[3.1.1 First analysis 23](#_Toc146485354)

[Dataset Description 23](#_Toc146485355)

[Data Dictionary (EDA) 23](#_Toc146485356)

[Data Preprocessing 23](#_Toc146485357)

[Data Visualisation 24](#_Toc146485358)

[Feature Selection 26](#_Toc146485359)

[Correlation Heatmap. 26](#_Toc146485360)

[Data resampling 27](#_Toc146485361)

[Model Training and Evaluation 27](#_Toc146485362)

[Machine Learning Models 28](#_Toc146485363)

[3.1.2 Second analysis 29](#_Toc146485364)

[Dataset Description. 29](#_Toc146485365)

[Data Dictionary (EDA) 29](#_Toc146485366)

[Data Preprocessing 30](#_Toc146485367)

[Correlation Matrix 30](#_Toc146485368)

[Data Visualisation 31](#_Toc146485369)

[CHAPTER 4 33](#_Toc146485370)

[4.1 Findings and Results 33](#_Toc146485371)

[4.1.1 Results First Analysis 33](#_Toc146485372)

[Comparative of Results before applying SMOTHE. 33](#_Toc146485373)

[Comparative Results after Applying SMOTHE. 34](#_Toc146485374)

[KNeighbors model 36](#_Toc146485375)

[Support Vector (SVC) model 36](#_Toc146485376)

[GaussianNB model 37](#_Toc146485377)

[Decision Tree Model 38](#_Toc146485378)

[38](#_Toc146485379)

[Random Forest model 38](#_Toc146485380)

[XGBoost Model 39](#_Toc146485381)

[Light GBM Model 39](#_Toc146485382)

[Gradient Boosting Model 40](#_Toc146485383)

[AdaBoost model 40](#_Toc146485384)

[Logistic Regression 41](#_Toc146485385)

[Results Comparison 42](#_Toc146485386)

[4.1.2 Results Second Analysis 42](#_Toc146485387)

[Results after applying Under-sampling for class imbalance. 42](#_Toc146485388)

[Random Under sampling 42](#_Toc146485389)

[Random Oversampling 43](#_Toc146485390)

[Lasso 43](#_Toc146485391)

[Random Forest Model 44](#_Toc146485392)

[Gradient Boosting Model 44](#_Toc146485393)

[AdaBoost Model 45](#_Toc146485394)

[Logistic Regression Model 45](#_Toc146485395)

[Support Vector SVM Model 46](#_Toc146485396)

[KNeighbors KNN Model 46](#_Toc146485397)

[Naive Bayes Model 47](#_Toc146485398)

[Decision Tree Model 47](#_Toc146485399)

[XGBoost Model 48](#_Toc146485400)

[LightGBM 48](#_Toc146485401)

[Results Comparison 49](#_Toc146485402)

[CHAPTER 4 50](#_Toc146485403)

[4.1 Conclusions and Future Research 50](#_Toc146485404)

[4.2 References 52](#_Toc146485405)

CHAPTER 1

# 1.1 Introduction

Credit card fraud represents an important fact to be considered for financial institutions, because of the large amount of reports that banks receive from customers regarding unknown transactions, which could be mentioned online purchases, withdraws, or transactions to other’s accounts, etc. As it could be identified the use of technology and the internet in general is increasing dramatically. (A. Banarescho, 2015). This is an important factor to consider because people have access to buy online anytime, anywhere. At the same time, financial institutions need to be prepared and contemplate the risk of fraudulent transactions which could be immersed in daily transactions. This requires developing an updated effective model, being able to give solutions, and mainly detecting the fraudulent transaction in an immediate stage. (A. Patcha, J. M. Park, 2007). Therefore, it will be possible to help banks and cardholders decrease the risk of large amounts of losses effectively and safely.

Credit card fraud detection is a difficult challenge that banks, and credit card issuers are attempting to solve by adopting fraud detection systems. In today's financial system, rule-based technologies are frequently used for fraud detection. However, the advancement and development of machine learning algorithms allow banks and financial organisations to recognise an unusual scenario faster for large financial data sets, as machine learning requires as many entries as possible to learn and predict with better accuracy. Different approaches were considered in this thesis.

In the first analysis, a dataset from Datacamp is employed, which are 339607 transactions, which occurred from 01/01/2019 to 31/12 /2020. The second dataset is from Deloitte, starting on 19/3/2018, and finishing in 31 days. The dataset doesn’t contain a date column, it contains a step column, which represents the hour of time.

According to (H. Bodepudi, 2021). The three methods studied, Isolation Forest (IF), Local Outlier Factor (LOF), and One-Class SVM, to discover the highest-performing unsupervised algorithms. The author investigated the performance of supervised algorithms, K-Nearest Neighbour (KNN), Decision Tree (DT), Logistic Regression (LR), and Random Forest (RF), for credit card fraud detection, and concluded that RF had the best performance for detecting fraud in credit card fraud transactions.

(M. Ummul Safa, R. M. Ganga, 2019 and C. Navamani, M. Phil, 2018). Compared the performance of three classification methods used for credit card fraud detection: Naive Bayes, K-Nearest Neighbour, and Logistic Regression, and their results revealed that the LR approach performed better than the other two. Further research on closest neighbor algorithms for anomaly detection.

For this thesis, an extensive analysis of credit card transactions is employed to detect credit card fraud. To achieve a good performance, it is necessary to apply different models that help to detect fraudulent transactions, The models considered in the analysis are, KNeighborsClassifier, SVC, GaussianNB, DecisionTreeClassifier, RandomForestClassifier, XGBClassifier, LGBMClassifier, GradientBoostingClassifier, AdaBoostClassifier, LogisticRegression, RandomForestClassifier, Over-sampling, Under-sampling and SMOTE. This aims to identify which machine learning model performs the best according to the dataset applied, and the approaches that are crucial for each model.

# 1.2 Research Question

Movement restrictions imposed throughout the world to combat the COVID-19 epidemic resulted in unprecedented growth in e-commerce, creating an opportunity for opportunistic fraudsters. The pandemic's influence on e-commerce fraud efforts has been felt internationally. In 2021, three-quarters of internet firms reported a net increase in these assaults compared to before the sanitary crisis. (Statista 2023). According to the most recent report, the rise in losses due to online fraud worldwide is a 140% increase from 2021 to 2023. This covers an essential problem: How can data analytics solutions, particularly predictive models, be evaluated and optimised for Financial Institutions to successfully minimise losses from credit card fraud?

# 1.3 Relevance

Considering that every second a massive number of transactions are made, banks require efficient technologies to process the information on time, as for humans wouldn't be possible to process and analyse such an amount of data in a limited time, also to investigate customer behaviour and identify possible patterns related with credit cards fraud detection. (K. Muameleci, 2022). The primary purpose of this thesis is to examine credit card transaction datasets, and machine learning techniques will be used to forecast if the transaction is legitimate or fraudulent. Different models are used to determine which one performs the best; among the machine learning algorithms used are: Logistic Regression, Random Forest, AdaBoost, Gradient Boostingand, SVM, KNN, Naive Bayes, and XGBoost, with the results being presented. Several sampling procedures will be utilised to address the class imbalance problem.

# 1.4 Contribution

This thesis's major contribution is to offer a better understanding of the importance for financial institutions to identify credit card fraud transactions in the early stage of the transaction, since the fact that this has become an enormous problem around the world, the project aims to identify distinct patterns connected to fraudulent transactions to aid in their identification. According to the latest surveys and studies, consumer behaviour changed dramatically after Covid 19, because of the facility and access to purchase whenever it needed it. On the other hand, this represents an essential activity for banks, since March 2020, buying patterns have shifted significantly, including how credit cards are used for online buying. Almost efficiently, scammers identified new methods of exploiting this with clever schemes and frauds, causing a significant increase in fraud cases.

Numerous nations are currently confronting a crisis of credit card fraud. It has become an important cause of concern for many countries. 459,297 cases of fraud involving credit cards have been recorded up to the year 2020. (Ian Wright. 2022). As technology advances, fraudsters now use complex schemes to acquire sensitive personal information from cards and then use that data to seize control of existing accounts or create new accounts for fraudulent identities. Most scammers use emails known as phishing to get control of people's bank accounts by obtaining confidential and private information.

# 1.5 Objectives

1. Apply different sampling techniques to undertake the class imbalance problem for machine learning algorithms.
2. Evaluate relevant machine learning models to provide the best forecast for credit fraud detection.
3. Provide various approaches based on findings to improve the early identification of fraudulent activities in financial institutions, resulting in a reduction in payment fraud losses.

CHAPTER 2

# 2.1 Literature Review

# 2.1.1 Credit card fraud application and findings.

The use of credit cards is considered one of the most important actions for banking transactions, since the fact that they deliver high profits and customer loyalty, however, the decision to approve a credit card is high risk for banks, nevertheless, with the appropriate application and use of the customer information the risk can be mitigated. The growing number of new card applications and the enormous outstanding amount of credit card bills during the recent pandemic make this even more challenging for banks to identify patterns for credit card approval. At the same time because of the high use of online purchases in daily life. Banks are even more exposed to risky transactions because of the facility for customers to obtain a credit card without difficulty, which results in a significant risk for losses and monitoring of detection in fraud transactions.

The importance of efficiency and prompt responses to customers are considered crucial for the customer when reporting an unusual movement in the account; the usage of machine intelligence for automating the detection process and moderating this challenge is recommended, as well as the productivity of such automation may depend on the richness of the model competence.

In-depth data analysis and preparation for the right training will determine if the model performs accurately for credit card fraud detection, and the availability of showing confidence for the machine learning which in turn can enhance the detection efficacy of fraudulent transactions.

Banks obtain a large number of credit card transactions every minute. Several of them are not complete for a variety of reasons, such as large amounts, lower amounts in the account, or too many queries on an individual's record. Manually analysing these programs is time-consuming, and errors risk, which means losses for the institutions, fortunately, with the power of machine learning, this work can be automated, and almost every commercial bank does so nowadays. In this project, are going to be used machine learning techniques to create automated credit card fraud detection, much like actual financial institutions.

To effectively ensure the effect of credit risk detection in science and technology finance, a credit risk prediction algorithm based on cloud computing is presented. To predict, the logistic regression model and increase the risk prediction capacity employing financial indicators in science and technology credit are chosen as a model of variables. (Li, Guiping. 2022)

A deep learning and machine learning model of credit prediction is built using industry data and enterprise data from tens of thousands of small and medium-sized businesses via data set division, processing, and model integration. First, using two characteristic selection strategies, multiple subsets of the dataset are evaluated using a convolutional neural network as the coarse prediction. (Zhang, Lei. He, Jie. Zhao, Zihao. 2022).

Credit risk management has increased considerably during the last decades, in terms of knowledgeable papers and the availability of methods for measuring and managing credit risk (Altman and Saunders 1998).

Current trends in credit risk management advocate the use of parametric, non-parametric, and ensemble models for credit default prediction, which are suitable for analysing large sample size data and provide better ways to capture complex relationships from the data (Figini et al. 2017; Lessmann et al. 2015; Butaru et al. 2016; Alaka et al. 2017).

Millions of credit card transactions are done every second, and people are unable to analyse and process such massive amounts of data in order to analyse fraudsters' behavioural patterns. This is where credit card fraud detection utilising machine learning algorithms comes in handy. There are two sorts of credit card fraud: online and offline fraud. However, Fragoso et al. (2018)'s typical method of prediction does not provide a single optimum model for tackling classification, a restriction in data for various probable combinations of predictors. Breiman (1996), and the availability of several modelling methodologies makes selecting the appropriate model challenging (e.g., Hastie et al. 2009; Kuhn and Jhonson 2013; Chipman et al. 2010).

The model averaging technique (Graefe 2014; Bates and Granger 1969), is an approach that provides high discriminatory power and precision compared to other traditional statistical methods. (Granger and Ramanathan 1984; Hansen 2007; Nelder and Wedderburn 1972), is one way to address such a limitation. Even though the model is an average of successful ways for handling issues, experimental implementation of model-be near to methods is difficult owing to model parameterization. This paper tackles this issue by offering a model average approach for linearly combining a series of biased models based on correlated-variate model prediction. To avoid any criticism, the proposed model does not emphasize parametrization. (Higgs and Banner 2017). To apply the plan, it is considered a novel methodology based on the solution of a quadratic equation compelling challenge. The proposed technique is based on the idea that the best average model is the one that minimizes the covariance between the errors of the individual models (parametric models, non-parametric models, and mixed models).

The proposed model's robustness is evaluated using a variety of key performance measures, including a measure (H), the area under the receiver operating characteristic curve (AUC), the area under the convex hull (AUCH), minimum error rate (MER), and minimum cost weighted error rate (MWL. This allows it to analyse the results' predictive capabilities, discriminatory power, and stability. When compared to well-known models, the suggested model's findings show superior performance.

In theory, the findings produced from the presented notion on a financial institution's dataset may be generalised to other groups of organisations for credit risk assessment (chance of default), because practically all entities have a dataset with a class imbalance of default risk even if there is a difference in the set of explanatory variables for the different dataset.

Most classification algorithms, which can be broadly classified as machine learning and artificial intelligence systems, are frequently not used by financial institutions due to stricter regulatory Committee requirements that support the use of parametric models for a simple and clear interpretation of the results. Despite the regulatory preference for adopting the statistical framework. (Ewanchuk and Frei 2019), a growing body of evidence supports the employment of sophisticated models in credit risk assessment (Leo et al. 2019).

(Alaka 2017), gives a comprehensive assessment of tool selection for analysing bankruptcy prediction models and addresses more advanced models for credit risk calculation.

(Chakraborty and Joseph 2017) advocate the use of a machine learning model to detect financial distress using balance sheet information, and their study concludes that the machine learning model outperforms the logistic regression model, which is the preferred classical approach of financial institutions.

(Khandani et al. 2010), used state-of-the-art non-parametric machine learning models to predict consumer credit risk default by combining transaction and credit data. The research shows that machine learning techniques may increase risk prediction more than traditional statistical approaches and that any subsequent lender loss can significantly be improved.

(Albanesi and Vamossy 2019), used a deep learning strategy based on a neural network and gradient boosting for high-dimensional data to forecast customer risk default. The work outperforms logistic regression models in terms of performance and adaptability to the aggregate behaviour of default risk.

(Bacham and Zhao 2017), compared the performance of machine learning models to industry-developed algorithms such as Moody's proprietary algorithm and proposed a 2-3 percentage point improvement in machine learning model performance. Although credit-behavior-related factors boost the discriminating strength of the studied models, the approach is slightly challenging to associate with the underlying company characteristics in forecasting credit risk default.

In estimating the credit risk default of small-medium firms, (Fantazzini and Figini 2009) suggested a non-parametric technique based on random survival forests. The performance comparison of the proposed model with the traditional logistic regression model reveals a weak relationship of performance between training and testing samples, implying an over-fitting problem, which is primarily due to contrasting logistic regression testing sample performance better than their proposed random survival models. Several more research, including (Kruppa 2013; Yuan 2015, Barboza 2017; Ampountolas 2021, and Addo 2018); demonstrate that machine learning outperforms any other statistical technique for credit risk prediction.

The literature on the non-statistical model frequently argues that the discrepancy between the expectation of the averaged forecasts and truth is dependent on the bias of contributing models as well as their weights. The underlying assumption for statistical model averaging literature, however, is that there is no bias, therefore their contribution is frequently less interesting (Burnham and Anderson 2002). Reducing bias is frequently highlighted as the major motivation for model averaging in many of the literature publications, particularly those linked to process models (Solomon et al. 2007; Gibbons et al. 2008; and Dietze 2017).

Weights are quadratic in terms rather than linear due to the nature of predictions, since knowing completely the correct approach to calculating weights is essential. (Breiman 1997) adds several advantages to the model averaging technique. Apart from the inaccuracy of the estimate, obtaining a decent estimator for the optimal weight in the first place is an open problem, and there is no such closed solution accessible, even in the case of linear models (Liang et al. 2011). The literature generally supports parametric, non-parametric, and ensemble model-averaging methodologies. Model averaging appears to be of importance for reducing prediction error as well as better reflecting model selection uncertainty (Buckland 1997; Madigan and Raftery 1994).

(Claeskens 2016) assumed that estimated model weights are beneficial in general since they are bias-free and have identical prediction variance, but this does not indicate that calculated equal weights are preferable. This field of study, to the knowledge, might be expanded by offering numerous suggestions for selecting weights, and the methodological approach outlined in this work is an effort in this direction to improve model predictive performance.

It is critical to understand that machines are not born intelligent. In general, supervised learning algorithms are trained to be clever by employing information gained from previous data. As a result, the historical data and learning algorithms are likely to prejudice the machines. The bias might render a computer incapable of dealing with undesirable scenarios for which it has not previously been taught. A human, on the other hand, can deal with such a problem, either by its own abilities or by collaborating with others. (Mehrabi, N. 2019).

Existing machine learning approaches generally assist the decision-making process by predicting or recommending the output of an observation. However, it is quite often reported in the literature that the end-users are unreliable about the trustworthiness of such a recommendation. It may be more prevalent in sensitive areas like finance, healthcare, etc.

Over time banks build an extensive customer database that can be analysed to evaluate the bank’s performance and make strategic decisions based on customers’ experience behaviour. This is a process that is still improving to find better approaches and precise models for the real world, and this is the reason why banks are always working on their customer experience and adapting to changes and new trends. Not all customers behave similarly regarding financial actions; therefore, a different treatment should be given to those who meet certainly profitable, this is becoming a big challenge for banks, especially for the credibility that a new customer must build, proving consistency to the institution so that, can be considered for upcoming applications.

Credit card companies utilize rule-based systems and other tools to identify fraud. One method is to utilize sophisticated fraud detection software. The model examines the transactions and determines whether they are fraudulent or legitimate based on past knowledge. Another method used by credit card issuers is to look for patterns used by credit card holders, which means that if the cardholder always uses the card in the same way, but suddenly a transaction falls outside of the card holder's normal pattern, the credit card company investigates whether that transaction is valid or not.

# 2.1.2 Types of electronic frauds.

These are some cases of credit cards that are related to electronic fraud, either directly or indirectly.

1. Credit Card Fraud: Credit cards, both virtual and real, are used to purchase supplies and services.

Virtual cards are used to commit fraud online, typically through the internet or phone, by getting credit card information illegally. Physical cards are used to commit fraud offline; the attacker must take the credit card.

1. Bankruptcy Fraud: Using a credit card while absent; concealing him; or engaging in other activities that cheat his creditors. Because of its intricacy, this form of fraud is difficult to foresee. (L. Delamaire, and J. Pointon. 2009).
2. Computer intrusion: the act of pushing one's way into getting unauthorised access to information with the intent of subverting the protection and detection mechanism.
3. Theft fraud / Counterfeit fraud: theft fraud is the use of a credit card without the owner's authorization, which may be checked as soon as the owner reports it to his financial institution. While credit card fraud offers the greatest risk, simply the credit card's data is necessary. (K. Chaudhary and B. Mallick, 2012).
4. Telecommunications: the use of telecommunication services to perpetrate various sorts of fraud is constantly changing; businesses, communication service providers, and consumers are all victims of this fraud. (K. Chaudhary and B. Mallick, 2012).

# 2.1.3 Different techniques used by credit card fraudsters.

Some of the most common credit card theft schemes are detailed below:

1. Credit card fraud-generating software: this is a computer software that creates genuine credit card numbers as well as expiration dates. These generators provide a list of credit card account numbers based on a single account number. The model operates by utilising the mathematical Luhm method, which is used by card issuers to produce additional acceptable card number combinations. This allows the user to generate as many numbers as he wants in the shape of any credit card format (T. P. Bhatla 2013). Black hat hackers sell compromised credit card information to criminals via illicit websites. (J. Akhilomen. 2013).
2. Physically stolen credit card information: A fraudster steals the card and uses the information for illicit purposes. It is possibly the most difficult type of traditional credit card fraud to combat.
3. CC/CVV2 shopping website: Fraudsters utilise stolen credit card information obtained from an illicit website to purchase goods and services. (J. Akhilomen. 2013).
4. Site cloning and merchant sites: fraudsters clone a full site, including only the pages where the client made transactions. Because the page seems like one of the genuine sites, the customer feels they are dealing with the firm from whom they desire to acquire products and services. The cloned site receives this information and sends an email acknowledging receipt of the purchase, just like the original firm.

The cloned site receives this information and sends an email acknowledging receipt of the purchase, just like the original firm. The thieves have all the information they need to perpetrate credit card theft. (T. P. Bhatla 2013). While merchant sites provide low-cost services to users and ask them to fill out their personal information, a fraudster can obtain a large number of credit cards.

1. Key-loggers and sniffers: The fraudster harms the user's computer by sending infected spam emails and requesting that the user download free games and software; this automatically installs a key-logger program that logs all keyboard input made into the computer on a file with the sole purpose of retrieving personal information over a network. Most of the time, this software is sold or shared on the internet among frauds.

# 2.1.4 Difficulties of credit card fraud detection.

Several problems stated below must be solved in order to properly accomplish fraud detection solutions and best practice performance (S. Sorournejad, Z. and A. H. Monadjem 2016).

1. Overlapping data: whether fraudulent transactions appear to be real or genuine transactions appear to be fraudulent; this is a significant difficulty that can lead to incorrect model design. (S. Sorournejad, Z. and A. H. Monadjem 2016.) (S. Maes K. Tuyls and B. Manderick 2002). (L.P. Andreas and J.S Salvatore. 2000).
2. Inability to adapt: Classification algorithms have the issue of recognising new patterns of fraudulent or normal behaviour. Most supervised or unsupervised fraud detection systems are incapable of detecting fraud.
3. Specifying a parameter: A lot of parameters, including a pit-set by the user, are required in the fraud detection task, which might lead to problematic model performance. This parameter has varying relevance, which increases the model's complexity. (T. P. Bhatla 2013).
4. A lack of standard metrics: The need for standardising access to and comparing good and negative results of fraud detection systems cannot be overstated.
5. Overfitting: This occurs when the algorithm used in model development attempts to learn as much information from the training data set as possible, even minor fluctuations that do not represent the real situation. This resulted in low prediction accuracy. (T. P. Bhatla 2013).

Credit card fraud detection techniques.

Fraud detection techniques are classified into two broad categories: fraud analysis (misuse detection) and user behaviour analysis (anomaly detection). (S. Maes K. Tuyls and B. Manderick 2002).

In anomaly detection, typical user behaviour is utilised to create a normal profile of the user, which is then used to check for large deviations from the normal user profile, which are deemed fraudulent transactions. This is an unsupervised pattern based on user account profile behaviour because each user, as well as the fraudsters, has their unique profile behaviour.

Incoming transactions are compared to a categorised supervised model of a known fraudulent transaction, which is programmed into a pattern to detect genuine and fraudulent transactions. To determine if a transaction is real or fraudulent, historical data is utilised to develop a categorization model. (S. Maes K. Tuyls and B. Manderick 2002). Briefly describe some of the most recent credit card fraud detection systems.

A diagram of a model

Description automatically generated with low confidence

## Figure 1. Data processing

Dealing with Fraud Detection Approaches.

There are several issues to handle when dealing with fraud detection systems, but four key ones are typically addressed:

To begin, idea drift is an issue that arises when a model has been trained and has learned a specific pattern of the consumer or imposter's activity, but then the behaviour changes. That is, the model is not successfully dynamic and does not adapt as quickly as the behaviour changes. As a result, it is critical for the efficiently recognise and categorise fraudulent activities as well as valid transactions. Second, there is a skewed class distribution. One of the most crucial difficulties confronting FDS is the highly unbalanced data.

There are several techniques to solve this challenge, including data-level and algorithmic-level approaches. The vast volume of data and its high dimensionality make data mining and detection extremely difficult. As a result, data reduction technologies such as dimensionality and numerosity reduction are commonly used. Principal Component Analysis (PCA) is a popular method for reducing dimensionality. Finally, the difficulty of real-time detection demonstrates the need for the system to detect fraud early in order to stop it or take action quickly. Various strategies, including Very Fast Decision Tree (VFDT) and Self-Organization Map (SOM), have been used to improve real-time detection.

# 2.1.5 Supervised Machine Learning.

Supervised learning is the process of categorising a new data point in the presence of labelled data. In other words, it uses labelled data to train models for categorisation of original data sets. Labelled data, for example, implies that it is known which occurrences are anomalies in areas where classification algorithms are utilised for anomaly identification. KNearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) are some categorization algorithms used in fraud detection. One of the most basic machine learning classifiers is the KNN algorithm. A data point is categorised by its nearest neighbors in this technique. KNN was proposed as an efficient algorithm for credit card fraud detection and was offered as a precise way to reduce the number of false alerts and detect fraudulent transactions.

Daniel Nelson 2021 describes the approach of a discriminative algorithm for partitioning the data space for a given labelled data set by finding an ideal hyperplane (a decision boundary in binary case). The authors examined the usage of SVM as a credit card fraud detection approach in high dimensional data sets and determined that this algorithm produces better results when utilising small data sets.

# 2.1.6 Clustering Methods.

Suggests two clustering techniques: peer group analysis and breakpoint analysis. (R. Bolton, & D. Hand. 2002), Peer group analysis refers to earlier accounts that were acting similarly but suddenly began behaving noticeably differently; the system finds these accounts and flags them as suspicious. While breakpoint analysis takes a different technique, when a large quantity is transferred, an account might be identified as suspicious and examined.

Imbalanced Dataset

This section discusses numerous techniques for resolving the class imbalance problem. The techniques may be divided into three categories: resampling approaches, ensemble-based approaches, and cost-sensitive learning approaches. This thesis only addresses the resampling strategy and the ensemble-based approach, which will be discussed in detail in the next parts. Cost-sensitive learning considers misclassification costs. In the medical diagnosis of cancer, for example, the cost of misclassifying a malignancy is substantially higher than the cost of projecting that a healthy individual has cancer. As a result, by weighting the misclassification cost of the minority class more severely than that of the majority class, the model's true positive rate may be increased.

Resampling approach

Most prediction models perform poorly in the context of an uneven class distribution. As a result, some data preparation must be conducted prior to delivering data as input to the model. In the event of a class imbalance problem, such data pretreatment is carried out utilizing a data-level method known as resampling. There are three types of resampling methods: under-sampling, oversampling, and hybrid.

The majority class is decreased in the under-sampling procedure to balance the dataset. When the size of the dataset is large, eliminating the bulk of samples can considerably increase performance and decrease storage issues. The oversampling method is the inverse of the under-sampling approach. This strategy is effective with the minority population. It duplicates minority class observations to equalize the ratio of the majority and minority samples. Finally, for rebalancing, a hybrid method employs both under-sampling and oversampling techniques. In the next sections, we will go through some of the resampling methodologies.

SMOTE stands for Synthetic Minority Over-sampling Technique.

SMOTE is a famous approach for rebalancing datasets that was created by Chawla BCHK02. Rather than oversampling with replacement, it seeks to generate new minority class examples (synthetic instances) by interpolating between multiple nearby minority cases. As a result, it reduces the problem of training data overfitting. The nearest neighbors of minority cases are chosen at random depending on the degree of oversampling necessary.

SMOTE and Tomek Link removal combined.

SMOTE is an effective method for balancing class distributions. However, the minority class cluster may infiltrate the majority class space while spawning new synthetic minority cases. Providing such information to the model may result in overfitting. As a result, both the SMOTE and Tomek Link elimination procedures may be used to balance the class distribution. The original training dataset is oversampled using SMOTE in this method, and then Tomek Link removal is performed to the resultant dataset to produce a balanced dataset.

# 2.1.7 HMM (Hidden Markov Model).

The hidden Markov Model is a limited set of states, each with its own probability distribution. To administrate transition between these states, a set of probabilities known as transition probability is employed (A. Singh and D. Narayan. 2012). The central concept is to construct a multilayer model of model behaviour based on both HMM and enumerating approaches for anomaly detection (T. Lane 1997). This methodology (HMM) does not require a fraud signature and may successfully identify fraud based just on the credit card owner's spending behaviour. The HMM examines the cardholder's spending habits based on a threshold value of high (h), medium (m), or low (l). This threshold value is dynamically determined by the clustering algorithm of each cardholder's personal expenditure routine. The most important advantage of the HMM-based technique is that it considerably minimises the number of valid transactions (false positives) identified as suspicious by the fraud detection system.

A picture containing text, screenshot, diagram, font

Description automatically generated

## Figure 2. Classification of fraudulent transactions.

# 2.1.8 SVM stands for Support Vector Machine.

This technique is appropriate for detecting credit card fraud since it requires two classes: valid and fraudulent. SVM attempts to compute an optional hyper lane that separates the sample of the two classes (D. Meyer 2012). SVMs are supervised learning models that use learning algorithms to analyse and recognise patterns for classification and regression tasks (N. Cristianini and J. Shawe-Taylor. 2000). Kernel representation and margin optimisation are two critical components of SVM. The optimal kernel for any given problem is a massive research challenge; speed and size (large training set), which decreases the demand computational for testing poses a key restriction to SVM.

The Decision Tree.

The decision tree is a diagram that depicts the potential consequences of a set of connected selections. It is used to create an algorithm that accurately predicts the optimal option. A decision tree may also be used to develop automated prediction models with many applications in data mining, machine learning, and so on. This approach can consider an item observation to forecast the value of that element. The advantage of this strategy is that it adds additional choices to an existing tree, is simple to grasp, and requires no data preparation. However, this approach has drawbacks in that it might grow very complicated and verify each operation one by one for better accuracy; several trees are frequently employed simultaneously in the ensemble method. (W. Fan, M.Miller, S. Stolfo. 2001).

Algorithm Genetic.

Genetic algorithms are heuristic search and optimisation methods that are encouraged by natural selection and belong to the wider family of evolutionary algorithm methods; these evolutionary algorithms have the propensity to get better solutions as time progresses. The challenge of fraud detection is a classification problem; GA has been applied in credit card fraud detection to minimise the number of transactions incorrectly categorised. (E. Duman, and H. M. Ozcelik 2011). GA is effective in detecting credit card fraud due to the ease with which programming languages may be implemented. It does, however, have a memory limitation and is time demanding.

The Artificial Neural Network.

This simulates how the human brain operates in certain circumstances, in order to accomplish the operations of nodes known as neurons. Neurons are computing units that process incoming data and generate output data. (E. Ngai, Y. Hu., Y. Wong 2011). A neural network is an interconnected network of nodes that reflect the linking functions of the human brain. (S. Ghosh, and D. L. Reilly 1994). ANN are nonlinear statistical data modelling tools that may build supervised/unsupervised learning patterns by modelling the complicated link between input and output. ANN is a random function approximation tool that can be learned by viewing datasets. In ANN, the terms "training" and "recognition" are frequently used. In the ANN-supervised training approach, sample data from both fraudulent and non-fraudulent transactions are utilised to develop models in fraud detection systems. (T. Guo, and L. Gui-Yang 2008).

Recurrent and LSTM Neural Networks.

RNNs (recurrent neural networks) are a subset of supervised machine learning models. They are composed of a series of cells with hidden states and non-linear dynamics. RNNs are typically applied to time series data, such as voice recognition, unsupervised anomaly detection, and automatic translation. In economics, LSTM is used as an alternative to the ARIMA model to forecast time series data. (Malhotra P. 2015).

Because credit card transactional data is temporal in nature, RNNs should be used instead of other types such as fully connected or convolutional neural networks. This is one of the models that is more commonly used nowadays for financial risk prediction analysis. It is important to highlight that this model is widely used in economics for predicting, implying that it is used in conjunction with recognition and unsupervised learning.

Over the years, financial institutions have investigated, applied, and polished the detection of credit card fraud transactions, and as a result, most banks now provide an expedient service to applicants. However, detecting and trusting artificial intelligence when it comes to money risk remains tough. At the same time, the availability of banks to detect fraud transactions becoming increasingly important, particularly for those who prefer to go for online shopping, due to the ability to do it at any time. One method utilised in the research is to use feature selection on characteristics acquired from raw transactional data to compare models and decide which performs better when using machine learning.

In credit scoring difficulties, feature selection was employed. In general, feature selection is critical for applications such as knowledge discovery in databases. As a result, the study in this work is driven by the need to automatically assess fraudulent transactions to make risk judgments, as well as the usage of credit card scores to make critical financial security decisions. Banks may use such ratings to categorise consumers into "risk groups," which might aid in detecting possible bankruptcy early and blocking the customer's card in time to reduce losses.

Neuron architecture will be stimulated by the usage of Neural Networks (NN) for machine learning frameworks. These are shown to simulate the human brain's ability to recognise complicated relationships between inputs and outputs. Based on prior studies, the researchers discovered the potential for the study and its relevance to the issue since there are sufficient data resources that can be enhanced with the treatment of another predictive model. (Haykin SS. 2009).

The model averaging strategy is largely effective for minimising prediction errors, although it may not be applicable in all situations. This is because a few individual models in the pool of models do not contribute to the reduction in covariance and average bias. This may be countered by adopting a suitable or diversified weight estimation approach, which helps to adjust the excess variance from weaker models.

The conventional method recommends choosing the single best model, which overlooks model uncertainty caused by model structure and assumptions. As a result, depending on the single best model with certainty is not a smart idea because it may have negative implications. When based on model average procedures, the committee of varied models improves performance. (Figini et al. 2016; Figini and Giudici 2017).

# 2.2 Validity

The result's validity is determined by k-fold cross-validation, which shows the difference between the prediction and the actual classification of the labelled features, as well as the accuracy of the model-sum of true positive and true negative predictions divided by the sum of all predictions. These measures are applied to the separated training and testing data, and if metrics like accuracy, recall, and precision show no overfitting or underfitting behaviour, the model is considered to be fit. If this condition is not met, the preceding parts of the study must be assessed, with the potential of more research necessary.

Once the models have been validated in terms of no overfitting or underfitting, the best models will be compared so that a single best model can be chosen as a recommendation for financial institutions to identify the best approach for credit card fraud detection. This procedure is described in the Methodology and Results chapters.

# 2.3 Sampling Strategy

The application of stratified sampling in the context of credit card fraud detection involves a meticulous approach to ensure the accuracy and effectiveness of the models employed. In this research endeavour, for analysis was applied ten distinct machine learning models across two diverse datasets to develop various strategies for addressing the complex challenge of credit card fraud detection. Stratified sampling played a pivotal role in this process.

Initially, the population of credit card transactions was stratified based on critical characteristics, such as transaction amount, location, age, type of transaction, and other cardholder features. Consequently, random samples were extracted from each section, with the sample sizes being determined in proportion to the representation of each subgroup within the entire population. This meticulous stratification and sampling approach offered several advantages. It substantially reduced sampling variability compared to simpler random sampling methods, resulting in more accurate estimations of population parameters. Moreover, it allowed for the precise analysis of different transaction subsets, enhancing the ability to identify fraudulent patterns within diverse groups of transactions. By ensuring that the sample was highly representative of the entire population, the research findings became more generalizable and applicable to a broader context in the realm of credit card fraud detection. Stratified sampling emerged as a powerful tool in mitigating bias and achieving robust, reliable, and actionable insights in this critical domain. Its practical applications extend beyond credit card fraud detection, encompassing various fields like market research, political polling, healthcare studies, and social sciences, where researchers strive to attain accurate and comprehensive samples that mirror the diversity of the populations under investigation. In summary, stratified sampling stands as an invaluable technique in primary research, elevating the quality of findings and bolstering the validity and trustworthiness of research conclusions, especially when dealing with intricate and multifaceted challenges like credit card fraud detection.

# 2.4 Primary Research and Ethics.

The initial plan was to conduct in-depth interviews with Credit Card Fraud specialists working for banks or other financial organisations. To accomplish this, actions were taken, such as contacting professionals who work for organisations dealing with fraud transactions via email and in person, but the response was unsuccessful since the organisations are unable to permit external individuals to get in touch with them regarding this subject, for the reason of data protection and particular for the sensitive and delicate data which might have been associated.

The main goal of this project is to identify and analyse the best approach to dealing with fraud and financial institutions, which is a big challenge nowadays, because of the number of frauds and scams around the world, especially in online shopping. This enables banks to anticipate situations where banks need to provide a quick response to customers, which is the reason why the investigation takes place; to achieve real-time data.

Simultaneously, the Data Analysis gained after implementing Machine Learning models for the project will provide intriguing issues to explore with the implementations of different methodologies that assist in collecting a better knowledge of how to deal with real-world circumstances regarding credit card fraud detection.

It is important to point out that ethical issues must be considered into account while conducting interviews with specialists in the field of credit card fraud detection, because of The European Data Protection Board where details; in accordance with Article 70(1)(e) of Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons in relation to the processing of personal data and the free movement of such data, and repealing Directive 95/46/EC (hereinafter "GDPR"). These recommendations aim to encourage a coordinated application of data protection guidelines about the administering of credit card data within the European Economic Area (EEA), as well as to ensure standardised protection of data subjects' rights, in full compliance with the GDPR's fundamental data protection principles.

The strategies investigated and discussed in this section of the study are offered to contribute to the overall goal of identifying fraudulent credit card transactions. distinct models are employed in this study to detect fraudulent credit card transactions; it is important to note that two distinct analyses are utilised to achieve an improved understanding of the problem, and particularly to gain better insights for fraud detection. The main reason why the two analyses are separate is that the datasets are completely different and there is no possibility of merging them since they all contain very different features that would cause bias to be analysed, and different approaches are applied regarding the objective of each of them.

In conclusion, the primary research for this thesis is based on Stratified sampling, which is a significant strategy in primary research, improving the quality of findings and enhancing the validity and credibility of research conclusions, particularly when dealing with complex and varied issues such as credit card fraud detection.

CHAPTER 3

# 3.1 Methodology

# 3.1.1 First analysis

Dataset Description

The dataset applied is from Datacamp, which contains 339607 card transactions spread out over two years from 01-01-2019 to 31-12-2020. Contains an amount of 339607 rows and 15 columns.

Data Dictionary (EDA)

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| transdatetrans\_time | Date and Time when the transaction was made. |
| merchant | Name of the merchant related to the location. |
| Category | Category of the transactions. |
| amt | Amount for each transaction. |
| City | The credit card holder's city location. |
| State | The credit card holder's state location. |
| Lat | Purchase latitude location. |
| Long | Purchase longitude location. |
| city\_pop | Credit card holder's city population. |
| job | Credit card holder’s Job. |
| Dob | Credit card holder’s date of Birth. |
| trans\_num | Transaction unique number. |
| merch\_lat | Merchant's latitude location. |
| merch\_long | Merchant's longitude location. |
| is\_fraud | Whether the transaction is Fraud (1) or Legit (0) |

*Table 1. Data dictionary first analysis.*

There is no missing data and after pre-processing the date out of the 337825 transactions, only 1782 transactions are fraudulent, making the dataset highly imbalanced. Further, the principal features for the analysis are presented in the features ‘trans\_date\_trans\_time’ of the transactions and ‘amt’ which shows the amount of each transaction made. The dataset was released to the public with the transactions labelled as fraud or not a fraud.

Data Preprocessing

In the initial stages of training machine learning algorithms to analyse statistical patterns and correlations, the collection and organization of sample data into datasets are foundational steps essential for achieving optimal results. Before constructing machine learning models, comprehensive dataset preparation is undertaken, encompassing various critical steps. These steps involve feature exploration to identify principal components and those requiring more in-depth investigation, transforming columns such as encoding categorical values into numerical formats and dissecting timestamp data into year, month, day, hour, minute, and second components to uncover nuanced patterns in credit card fraud transactions. Additionally, data is meticulously processed by verifying the absence of missing values and the elimination of duplicate information to ensure clean and reliable inputs. Given that the 'amount' variable exhibits non-normal distribution characteristics, feature scaling becomes imperative to address skewness, particularly crucial in fraud detection, as skewed features can significantly impact the performance of statistical models. This may involve data transformations or preprocessing techniques to enhance feature suitability for subsequent modelling endeavours.

Data Visualisation

This graph shows that the dataset is significantly skewed; only 1782 transactions are classed as fraudulent, whereas 337825 transactions are classified as legitimate.

## A blue square with white text Description automatically generatedFigure 3. Credit Card Fraud and Legit Transactions.

A graph of a graph

Description automatically generated with medium confidenceThis graph depicts the hours when legitimate and fraudulent transactions occurred. It can be seen that legitimate transactions are mostly made after 11 a.m., whereas fraudulent transactions occur before 5 a.m., providing a unique insight into hour transactions.

## Figure 4. Credit Card Fraud and Legit Transactions per hour.

A graph of a graph of a graph

Description automatically generated with medium confidenceIn terms of category, the largest representation is "shopping\_net," which represents all online shopping, and the second highest is "food\_dining," which represents food transactions that may also be paid for online. According to this graph, there may be a correlation between the categories of fraudulent transactions.

## Figure 5. Credit Card Fraud and Legit Transactions per category.

Based on the feature state, can be concluded that the number of fraudulent transactions follows the same pattern as genuine transactions. This suggests that the number of fraudulent transactions is more closely tied to the total number of transactions and that there is no meaningful association between fraudulent transactions and the column state.

A graph of a graph

Description automatically generated with medium confidence*Figure 6. Credit Card Fraud and Legit Transactions per state.*

A graph with different colored bars

Description automatically generatedDifferent behaviour may now be identified individually for fraudulent operations, which aids in determining whether there is a link or significant insights into the values of the characteristics and fraudulent transactions. According to Category, fraud transactions are more prevalent in the'shopping\_net' and 'home' categories. The state with the most fraudulent transactions is indicated by the number two; however, based on the previous graph, it does not represent a relationship with fraudulent transactions; rather, it is related to the number of transactions made in general per state, as shown in the same graph as legitimate transactions.The behaviour for hours provides a crucial insight, indicating that the majority of fraud transactions occur between the hours of 22 p.m. and 3 a.m.

## Figure 7. Credit Card Fraud Transactions per categories.

Almost a quarter of all fraudulent transactions were for less than $50, which is an important indication for fraud prediction. More than 300 transactions were carried out for amounts ranging from 800 to 1,000. According to the graph, the transactions between $400 and 600 were the fewest. A single transaction can have a maximum value of 1,371.81 and a minimum value of 1.78.

A graph showing a number of transactions

Description automatically generated

## Figure 8. Credit Card Fraud Transactions per categories.

Feature Selection

Each of the characteristics obtained from the dataset may not be useful in developing a machine learning model to make the required prediction. Some of the features may increase forecast accuracy. As a result, feature correlation plays an important role in developing a stronger machine learning model. High correlation features are more likely to be linearly related and have virtually the same influence on the dependent variable. As a result, when two characteristics have a strong correlation, we can drop one of them. The correlation heatmap of the original dataset and resampled dataset.

A graph showing a number of different colored bars

Description automatically generated with medium confidence

## Figure 9. Feature importance scores.

Correlation Heatmap.

The first analysis employed necessary techniques to identify the relevant features and its correlation to the target variable, such as Correlation heatmap, Feature Importance Score, etc. After applying these different techniques for feature engineering a certain improvement in the model is obtained, it shows more confidence to proceed with the machine learning models.

Different techniques were attempted in order to find the connection between the characteristics and also the relevance of including all of the features for the model; to detect this, LASSO selection was used; nevertheless, the model did not improve with the strategy. At the same time, amt and log\_amt have a good correlation; nevertheless, based on the findings, the models performed better when both characteristics were included; so, both will be used for modelling. The attributes having the best link to the target variable, according to the correlation matrix, are amt, log\_amt, lat, and hour.

A screenshot of a computer

Description automatically generated

## Figure 10. Correlation Matrix-Analisis 1

Data resampling

As previously stated, the dataset is strongly imbalanced. The number of valid transactions exceeds the number of fraudulent ones. In this situation, if it utilises this dataset to train the model, the model will be biassed towards valid transactions, resulting in poor model performance when evaluated on unknown data. For this study is used resampling approaches such as random under sampling, random oversampling, SMOTE, Tomek links removal, and a combination of SMOTE and Tomek links removal to solve this problem. To balance the training data, as well are used various resampling strategies individually.

Model Training and Evaluation

This study, conduct experiments with both supervised and unsupervised machine learning model to classify fraudulent transactions. It also discusses the process of model creation and selecting the values of the hyperparameters for the best model. After placing the hyperparameters into each model and tweaking them for each prediction model, and then gave the resampled training set to each model as training data. As a result, the models discovered new patterns in the resampled training data. The model's performance is then tested using the test set, which is used previously segregated while partitioning the entire dataset.

Machine Learning Models

Accuracy score, F1 Score, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R2 Score, and Classification report were used for all the following models, and the findings will be described in the next phase.

**KNeighbors:** This model is applied because helps for tasks such as classification, regression, and anomaly detection. It is suitable for both binary and multiclass classification problems, since the model requires a minimum assumption about the data distribution, it is important to preprocess the data appropriately.

**Support Vector (SVC):** This model is a powerful algorithm for the purpose of fraud detection, as it is explained previously the dataset contains different features with high dimensional data, that SVC can effectively handle and analyse. At the same time is a good model since fraudulent activities may not present linear patterns that could be challenging with linear models. SVC is suitable for imbalanced datasets, and it will be a balance for the SMOTE technique applied previously.

**GaussianNB:** This is a simple and computationally effective algorithm that is suitable to train and make predictions. It is ideal for datasets with continuous or quantitative characteristics as the following dataset. Transaction amounts, timestamps, and other continuous variables are used in fraud detection to determine irregularities. The model is employed also because of its probability of confidence associated with each prediction, this is an important factor for decision-making and risk assessment.

**Decision Tree:** The application of this model is essential for credit card fraud detection, because of its easy interpretability which helps to understand why is making the decision. The feature importance also benefits because provides insight to identify which features are the most relevant for making fraud detection decisions. Also, can detect non-linear relations between features and the classification labels. This model is also adaptable, so be customised with specific business rules, which makes it suitable for real-time scoring, so can make predictions quickly, as it is needed for fraud detection to make swift decisions on incoming transactions.

**Random Forest:** This is a powerful and wide machine learning algorithm that generally delivers high predictive accuracy, as well as ensemble learning methods that combine the decision of multiple decision trees. It can capture complex relationships between features and the classification of labels. It is approachable with the SMOTE technique and furthermore is relatively robust to outliers that could be present in the dataset.

**Extreme Gradient Boosting XGBoost:** This is an essential model for credit card fraud detection since provides feature importance scores, indicating the significance of each feature in making predictions. It normally predicts with exceptional accuracy and is a gradient-boosting ensemble method that combines the prediction of multiple models generally Decision Tree. Because of its capacity to handle complicated, high-dimensional, and unbalanced datasets, its resistance to noise and outliers, and its remarkable prediction accuracy, XGBoost is a great tool for credit card fraud detection. When properly configured and optimised, XGBoost may assist organisations in detecting and mitigating fraudulent actions while minimising the impact on legitimate clients.

**Light GBM:** This model is recognised for its high prediction accuracy. It can detect nuanced fraud tendencies while minimising false positives by capturing complicated interactions between features and class labels.It is based on gradient boosting, which is an ensemble approach for combining the predictions of numerous low learners. This ensemble technique eliminates overfitting while improving model generalisation, resulting in better fraud detection performance.

**Gradient Boosting:** This model is an effective approach for detecting credit card fraud because of its capacity to handle complicated, high-dimensional, and unbalanced datasets, resistance to noise and outliers, and remarkable prediction accuracy. Gradient Boosting, when designed and calibrated appropriately, may assist organisations in successfully detecting and mitigating fraudulent actions while minimising disturbances to real consumers.

**Adaptive Boosting (AdaBoost):** Because of its capacity to handle complicated, high-dimensional, and unbalanced datasets, its resistance to noise and outliers, and its remarkable prediction accuracy, is a valuable tool for credit card fraud detection. This model requires proper designing and engineering, therefore can assist organisations in successfully detecting and mitigating fraudulent actions while minimising disturbances to legitimate consumers.

**Logistic Regression:** In this model, a feature is assigned to a coefficient that indicates the direction and intensity of its effect on the expected outcome. This makes it simple to determine which attributes to fraud detection and how. It is a straightforward and linear model. This is computationally efficient and only requires a small number of parameters to be taught. This facilitates its use and is beneficial when working with real-time or near-real-time fraud detection systems that must make instant choices.

3.1.2 Second analysis

Dataset Description.

The dataset applied is from Deloitte and contains 6362620 card transactions spread out over 31 days, the dataset does not contain a date column, however, contains a step column, which represents the hour for the transaction from 1 to 743, Contains an amount of 6362620 rows and 11 columns.

Data Dictionary (EDA)

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| step | Date and time of the transaction, ranging from 1 to 743, representing 31 days considered for data collecting 24 hours a day. |
| type | Following type of transaction presented: 'PAYMENT', 'TRANSFER', 'CASH\_OUT', 'DEBIT', 'CASH\_IN' |
| amount | Amount for each transaction. |
| nameOrig | Is a unique alphanumeric number assigned to each transaction that refers to the origin of the transaction. |
| oldbalanceOrg | Represents the sender's balance prior to the transaction. |
| newbalanceOrig | Represent the sender's balance after the transaction. |
| nameDest | Is a unique alphanumeric number assigned to each transaction that refers to the transaction's destination. |
| oldbalanceDest | Represents the sender's balance before the transaction. |
| newbalanceDest | Represents the sender's balance after the transaction. |
| isFraud | Whether the transaction is Fraud (1) or Legit (0) |
| isFlaggedFraud | Whether the transaction is Fraud (1) or Legit (0) |

*Table 2. Data dictionary second analysis.*

There is no missing data and after applying data processing of the 6354407 transactions, only 8213 transactions are fraudulent, making the dataset highly imbalanced. The dataset does not contain a variety of features as the previous dataset, these features are more related to the bank account holder rather than credit card transaction. The most relevant column is ‘amount’ which shows the quantity of each transaction made. The dataset was released to the public with the transactions labelled as fraud or not a fraud.

Data Preprocessing

To train machine learning algorithms to assess statistical patterns and correlations, sample data must be gathered and saved in datasets; this phase is the start of the discoveries, but it is critical to produce a satisfactory outcome. Prior to developing machine learning models. The following approaches are included in the second dataset:

In the initial phase of data preparation, we explored the dataset's features, pinpointing key variables while recognizing those requiring further investigation. We also identified columns necessitating transformation, particularly converting categorical data into numerical formats to facilitate future machine learning applications. Moreover, we converted a specific column into day and hour components, enabling the detection of patterns related to credit card fraud transactions, building upon prior findings emphasizing the significance of the "hour" feature. Additionally, we rigorously examined the dataset for any missing values and duplicates to ensure data integrity and prevent noise in subsequent analyses. To avoid over-fitting, we assessed feature similarities and ultimately transformed all data types into integers, aligning the dataset for compatibility with machine learning algorithms.

Correlation Matrix

Based on the correlation heatmap, the column hour has a negative correlation with the target variable; however, previous analysis revealed that hour is a relevant feature for credit card fraud detection; therefore, the columns will not be removed until a deeper comprehension of the values has been established.

The columns oldbalanceOrg and newbalanceOrig, as well as oldbalanceDest and newbalanceDest, show a high scale of correlation; these characteristics are more likely to be linearly connected and have approximately the same impact on the dependent variable. Therefore, if two features in the dataset are highly correlated, the outcomes of the machine learning model could be affected, especially if the correlation is so strong that it creates multicollinearity, which means that two or more independent variables in a regression model are significantly connected with each other. As a result, it is possible to drop one of the features.

After removing the characteristics that are highly correlated with one another, the correlation heatmap looks like this:

A graph of a heatmap

Description automatically generated

## Figure 11. Correlation Heatmap after dropping unnecessary Features - Analysis 2

Following the use of under-sampling to correct class imbalances, methods for training samples for the minority class were created. Because the correlation between the features did not change much, all the columns will be utilised for modelling. Even though the Lasso approach was already tested, the model did not improve after maintaining just the relevant characteristics that the model recommended.

Data Visualisation

According to the graphs presented, fraudulent transactions represent only 1% of the total dataset, indicating that the dataset is highly imbalanced, and if the models are processed in this manner, the results will rise to predict the entire sample as legitimate transactions because the models will not receive enough data to train how to predict the fraudulent transactions; as a result, the under-sampling technique is used to process and prepare the information.

A blue rectangle with numbers and a white background

Description automatically generatedA purple circle with text

Description automatically generated

## Figures 12, 13. Credit card fraud transactions vs. legit transactions - Analysis 2

A comparison of a bar graph

Description automatically generatedCan be identified that fraudulent transactions are represented by the values 1 and 4, which means that are just: TRANSFER and CASH IN, the rest of the features don’t present any operation for the column type.

## Figure 14. Credit card transactions by type - Analysis 2

A graph of a graph of a graph

Description automatically generated with medium confidenceAs can be seen, fraudulent transactions are not presented at a certain time; they appear to be more prevalent from 5 am to 10 am; however, it can detect a significant decrease at 3 am, when it is a high occurrence according to analysis 1. This demonstrates that there is no connection between hours and fraudulent transactions.

## 

## Figure 15. Credit card transactions by hour - Analysis 2

CHAPTER 4

# 4.1 Findings and Results

# 4.1.1 Results First Analysis

Comparative of Results before applying SMOTHE.

Prior to using SMOTE, the results of machine learning algorithms obtained were very poor in detecting fraudulent transactions. This is because only a minority of the training samples were labelled as fraud transactions, which represents a huge challenge for the algorithm to identify patterns in fraud transactions because the majority of the transactions are legit, causing the algorithm to lean towards predicting it as a legit transaction.

In machine learning, the classification issue may be defined as the challenge of predicting the class label of a given data point. Fraud detection, for example, can be identified as a classification challenge. The purpose of this situation is to anticipate whether a particular transaction is fraudulent or real. There are three types of classification in general: binary classification, where there are two output labels (e.g., classifying a transaction that may be fraudulent or genuine), multi-class classification, where there are more than two output labels (e.g., classifying a set of images of animals that may be cat, dog, or cow). And multi-label classification, in which the data samples are not mutually exclusive, and each data sample is assigned a set of target labels (for example, categorising a crab based on sex and colour, with output labels that can be male/female and red/black). This thesis is concerned with the binary classification issue, in which the output label is either normal or fraudulent. (Nathalie Japkowicz and Shaju Stephen, 2002)

Most real-world applications have an uneven class distribution, in which the count of one class label far outnumbers the count of another. The fraud detection job is a good example of a class imbalance problem since the number of fraud class labels is quite low in comparison to the number of regular class labels. In the face of an uneven class distribution, most machine learning methods perform poorly (i.e., the predictive model tends to categorise the minority case as the majority example). As a result, several problems and challenges may arise. Here are the results of fraud prediction of the model’s metrics performance before employing SMOTE, for oversampling the fraud transactions:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Models** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **AUC Score** |
| KNeighbors | 0.9962 | 0.86 | 0.36 | 0.5105 | 0.7939 |
| Support Vector (SVC) | 0.9963 | 0.92 | 0.36 | 0.5166 |  |
| GaussianNB | 0.9877 | 0.21 | 0.47 | 0.2944 |  |
| Decision Tree | 0.9974 | 0.75 | 0.79 | 0.7682 | 0.6385 |
| Random Forest: | 0.9981 | 0.94 | 0.70 | 0.8025 | 0.8612 |
| XGBoost | 0.9986 | 0.95 | 0.79 | 0.8647 | 0.9628 |
| Light GBM | 0.9958 | 0.60 | 0.71 | 0.6491 |  |
| Gradient Boosting | 0.9958 | 0.74 | 0.35 | 0.4788 | 0.9467 |
| AdaBoost | 0.9957 | 0.72 | 0.36 | 0.4791 |  |
| Logistic Regression | 0.9945 | 0.0 | 0.0 | 0.0 | 0.4963 |

*Table 3. Models Comparison (Before SMOTE)*

A graph with different colored bars

Description automatically generated

## Figure 16. Model Comparison Results (Before SMOTE)

According to the classification report, all the models did quite well in terms of accuracy, which is a critical parameter; nevertheless, the data reported for Recall and F1 Score do not show the same score. According to Recall, the model with the poorest performance is Logistic Regression, which indicates that the model could not identify even one of the real positive cases, earning a score of 0%. A poor recall shows that the model is missing a large number of positive examples. Following that, the models with the highest Recall scores are Decision Tree and XGBoost, both of which achieve 79%. What this means is that 290 positives were acquired from a total of 367 positive occurrences, with 77 of the positive features identified as negative.

Following this finding, it can be demonstrated that a high accuracy does not imply that the model worked optimally. To determine model performance, it is important to investigate a combination of the entire report and all metrics.

Comparative Results after Applying SMOTHE.

After applying SMOTE, for oversampling fraudulent transactions, the model performance improved much better regarding Recall and Precision, which helped with the challenge of the minority of fraudulent transactions in the dataset, without SMORE, all the models achieved a very high accuracy but performed very poorly in identifying fraudulent transactions, which is the principal objective. SMOTE helped address the class imbalance issue by oversampling the minority class and led to a model with a more balanced trade-off between precision and recall. When dealing with imbalanced datasets and tasks like fraud detection, it’s often more important to focus on metrics like F1 score, precision, and recall rather than accuracy.

The reported accuracy indicates the percentage of the model correctly classifying the transactions. However, it's essential to search deeper into the performance metrics to understand the model's behaviour better:

1. True Positives (TP): The model correctly classified fraudulent transactions as fraudulent.
2. False Positives (FP): The model incorrectly classified legitimate transactions as fraudulent.
3. True Negatives (TN): The model correctly classified legitimate transactions as legitimate.
4. False Negatives (FN): The model incorrectly classified fraudulent transactions as legitimate.

**Additional performance metrics:**

**Precision:** Precision is crucial for credit card fraud detection because it assesses the accuracy of positive predictions made by the model. It indicates the model's ability to correctly identify and flag potentially fraudulent transactions, minimizing the number of false alarms that could inconvenience customers. High precision ensures that when the model raises an alert, it is highly likely to be a true case of fraud, thus improving the overall effectiveness of fraud detection systems.

**Recall (Sensitivity):** Recall is equally important in fraud detection as it measures how well the model captures all actual instances of fraud. A high recall value signifies that the model is proficient at identifying most, if not all, fraudulent transactions, reducing the risk of undetected fraud cases. This is crucial for protecting both customers and the financial institution from substantial losses.

**F1 Score:** The F1 score, being a balanced measure of precision and recall, is essential for credit card fraud detection as it strikes a harmonious balance between minimizing false positives and false negatives. Achieving a high F1 score ensures that the model maintains a strong balance between precision and recall, optimizing its performance in accurately identifying and flagging fraudulent activities while minimizing errors.

**Specificity:** Specificity is relevant in fraud detection because it measures how well the model identifies true negatives, i.e., legitimate transactions. A high specificity indicates that the model effectively recognizes and approves genuine transactions, reducing the chances of valid transactions being erroneously declined or flagged as fraudulent.

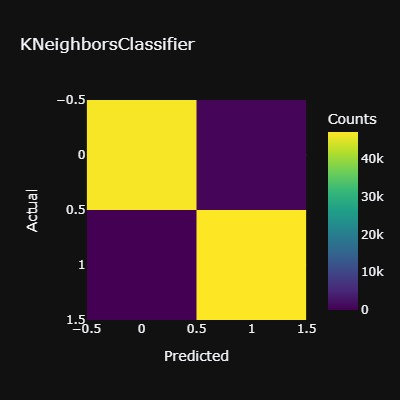
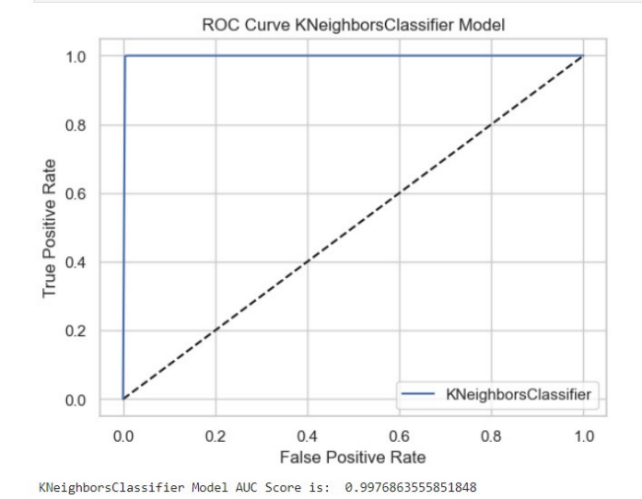
**False Positive Rate (FPR):** FPR is significant for credit card fraud detection as it quantifies the proportion of actual non-fraudulent transactions incorrectly classified as fraudulent. Lowering the FPR is essential to prevent inconveniencing customers with false alarms while maintaining a robust fraud detection system's accuracy. In fraud detection, achieving a high recall (capturing most fraudulent cases) while maintaining reasonable precision (avoiding too many false positives) is often crucial, as missing fraudulent transactions can be costly. It's also essential to consider the specific requirements and constraints of the application when evaluating and fine-tuning the model.

The following are the outcomes of including SMORE into the models:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Models** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **AUC Score** |
| KNeighbors | 0.99 | 0.99 | 1.00 | 0.99 | 0.9976 |
| Support Vector (SVC) | 0.98 | 0.97 | 0.98 | 0.98 |  |
| GaussianNB | 0.85 | 0.95 | 0.73 | 0.83 | 0.8491 |
| Decision Tree | 1.00 | 0.99 | 1.00 | 1.00 | 0.9950 |
| Random Forest | 1.00 | 1.00 | 1.00 | 1.00 | 0.9999 |
| XGBoost | 1.00 | 1.00 | 1.00 | 1.00 | 0.9999 |
| Light GBM | 0.99 | 0.99 | 0.99 | 0.99 | 0.9993 |
| Gradient Boosting | 0.94 | 0.96 | 0.93 | 0.94 | 0.9877 |
| AdaBoost | 0.91 | 0.92 | 0.90 | 0.91 | 0.9738 |
| Logistic Regression | 0.85 | 0.94 | 0.75 | 0.84 | 0.8917 |

*Table 4. Models Comparison (After SMOTE)*

KNeighbors model



## 

## Figure 17. KNeighbors Model performance First analysis

The model's accuracy is 0.99, indicating a very good performance. The model identified all of the fraudulent transactions, a total of 47154, but from the 47444 expected legit transactions, it classified 46733 as legits, implying that 711 of the legit transactions are classified as fraudulent rather than legitimate. This is an excellent result the model performed perfectly with just a few misidentified samples against the total amount.

Support Vector (SVC) model

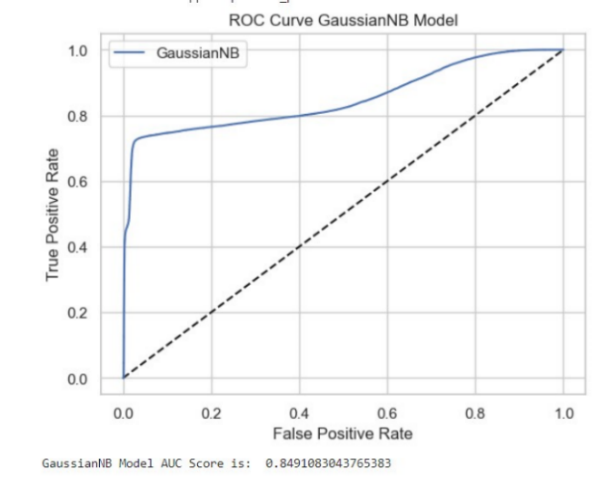
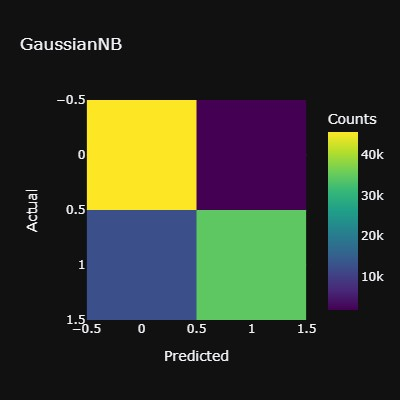
A screenshot of a graph

Description automatically generated

## Figure 18. SVCModel performance First analysis.

The model's accuracy is 0.98, indicating a very good performance. The model identified 46413 as fraudulent transactions, from a total of 47154, in the other hand from the total of 47444 expected legit transactions, it categorized 46107 as legit, implying that 1137 of the legit transactions the model predicted as fraudulent. Also, 741 of the fraudulent transactions are classified as legit rather than fraudulent. This is a very good result for a powerful learning algorithm that is known for its ability to handle complex classification tasks but also requires more effort in terms of hyperparameter tuning and kernel selection.

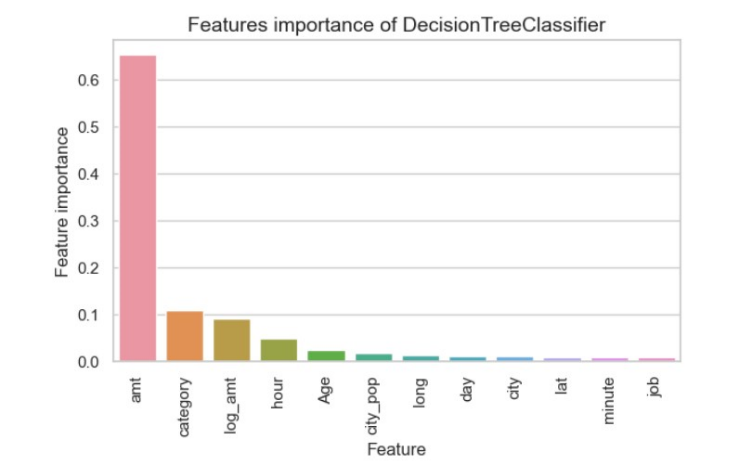
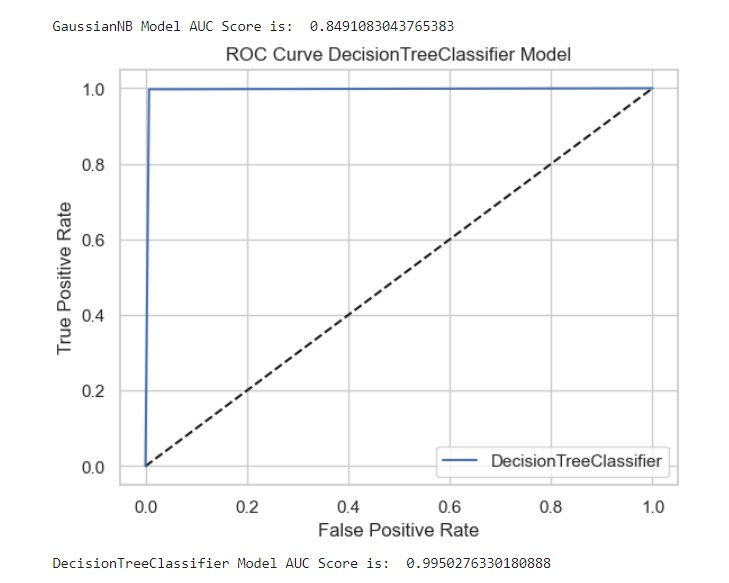
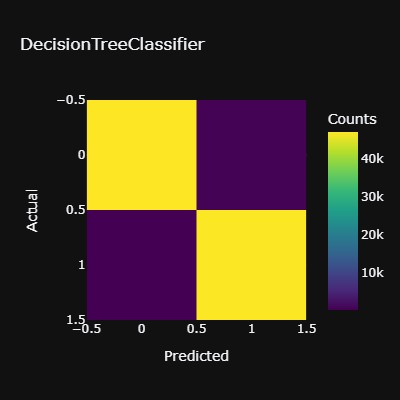
GaussianNB model



## Figure 19. GaussianNB Model performance First analysis.

The model's stated accuracy is 0.85, indicating satisfactory performance. The model identified 34474 fraudulent transactions out of a total of 47154, while projecting 45610 legitimate transactions out of a total of 47444 expected valid transactions, implying that the model misidentified 1834 legitimate transactions. Furthermore, in 12680 of the fraudulent transactions, the model was categorised as valid rather than fraudulent. Overall, the model performed well; nevertheless, the number of false positives would result in significant loss for a financial institution, and so would not be appropriate for the main purpose.

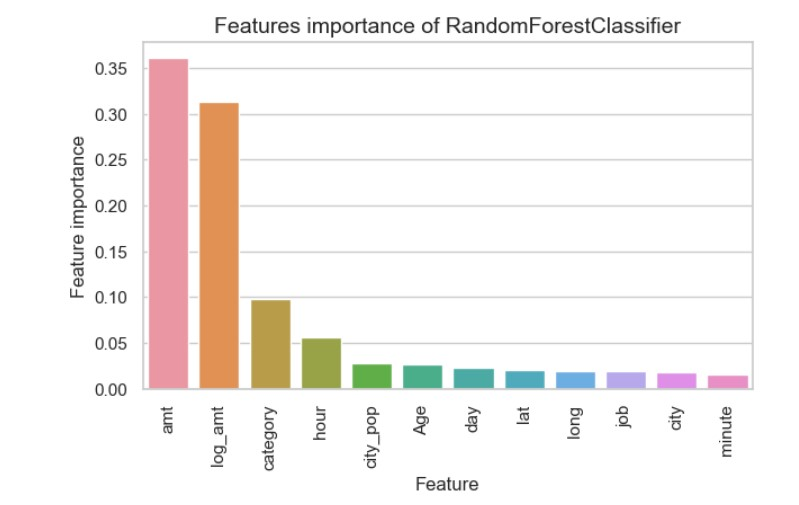
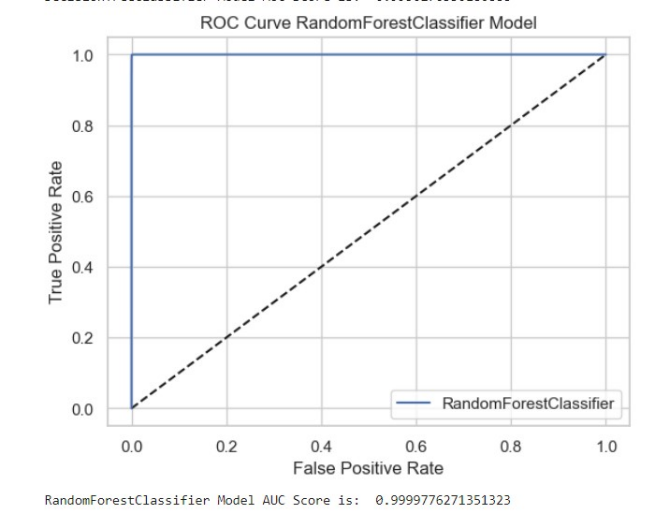
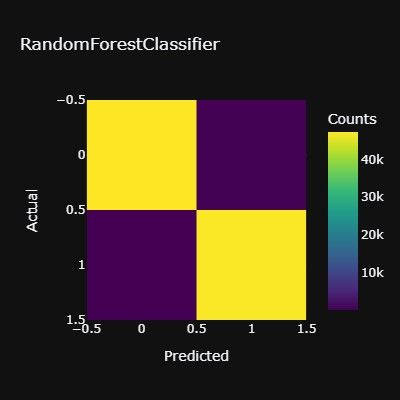
Decision Tree Model



## Figure 20. Decision Tree Model Performance First analysis.

The reported accuracy of the model is 1.00, signifying remarkable performance. The model recognised 47021 fraudulent transactions out of a total of 47154, while projected 47106 genuine transactions out of a total of 47444 predicted valid transactions, meaning that 338 legit transactions were misdiagnosed. Furthermore, the model was classified as legitimate rather than fraudulent in 133 of the fraudulent transactions. Overall, the model performed admirably, with just a tiny number of transactions failing to identify correctly. The most essential factor for decision-making in the model is "amt," followed by category, log\_amt,hour, and age, with the remainder of the features showing little to no association.

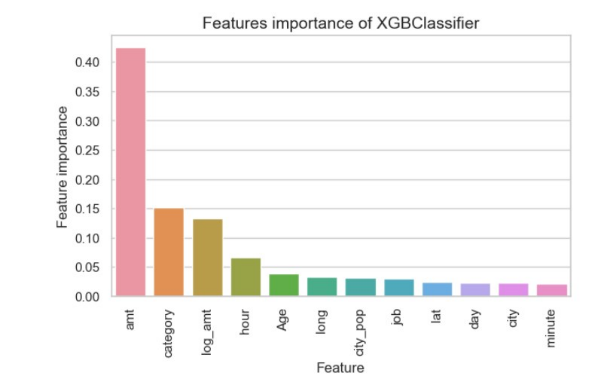
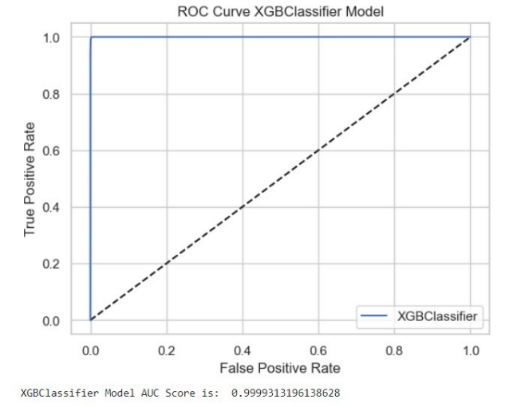
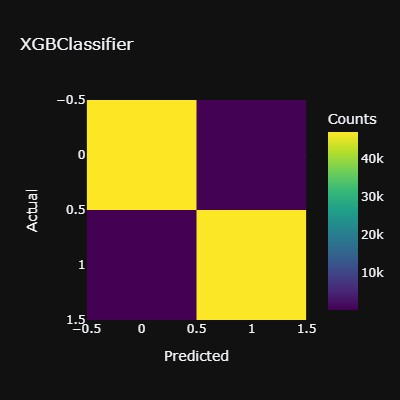
Random Forest model



## Figure 21. Random Forest Model Performance First analysis.

The model's stated accuracy is 1.00, indicating exceptional performance. The algorithm identified 47143 fraudulent transactions out of a total of 47154, while projecting 47328 real transactions out of 47444 anticipated valid transactions, implying that 116 legitimate transactions were misidentified. Furthermore, in 11 of the fraudulent transactions, the model was labelled as valid rather than fraudulent. Overall, the model worked remarkable, with just a small percentage of transactions failing to identify correctly. In terms of fraudulent transaction prediction, it is a very strong score to miss only 11 out of 47154. The model's most important decision-making factors are "amt" and "log\_amt," followed by category and hour, with the remaining characteristics exhibiting some relationship in general.

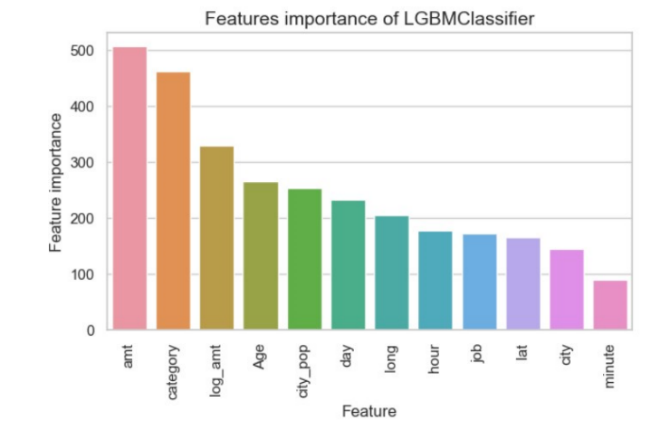
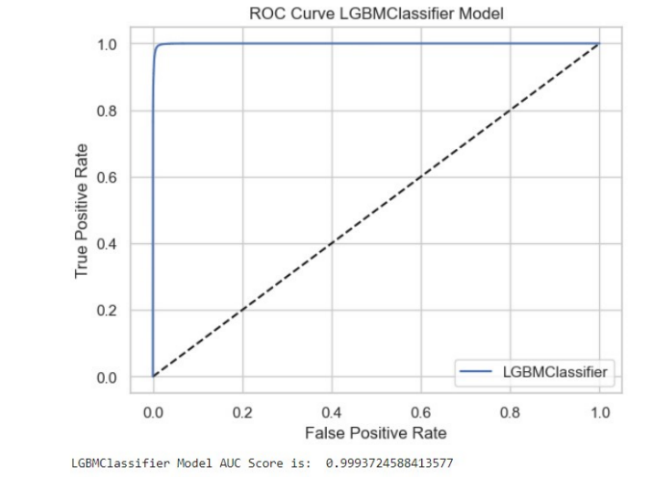
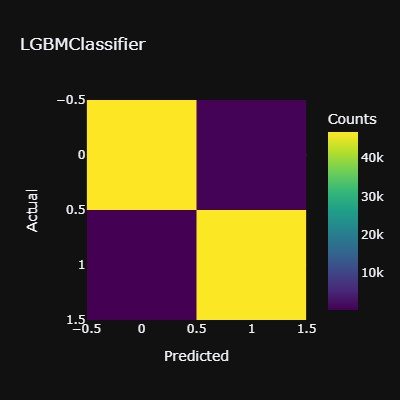
XGBoost Model



## Figure 22. XGBoost Model Performance First analysis.

The reported accuracy of the model is 1.00, signifying remarkable performance. The programme detected 47121 fraudulent transactions out of a total of 47154, while anticipating 47255 actual transactions out of 47444 expected legal transactions, meaning that 189 legitimate transactions were misdiagnosed. Furthermore, the model was identified as genuine rather than fraudulent in 33 of the fraudulent transactions. Overall, the model performed admirably, with just a tiny number of transactions failing to accurately identify. In terms of fraudulent transaction prediction, a score of 33 out of 47154 is quite good. "amt" is the most important decision-making component in the model, followed by category, log\_amt, and hour, with the remaining attributes demonstrating some association in general.

Light GBM Model



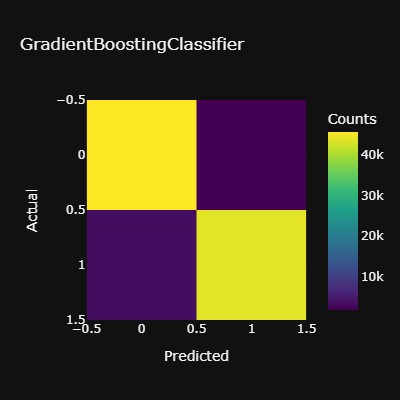
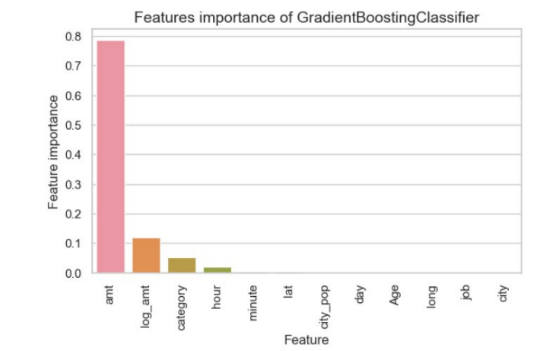
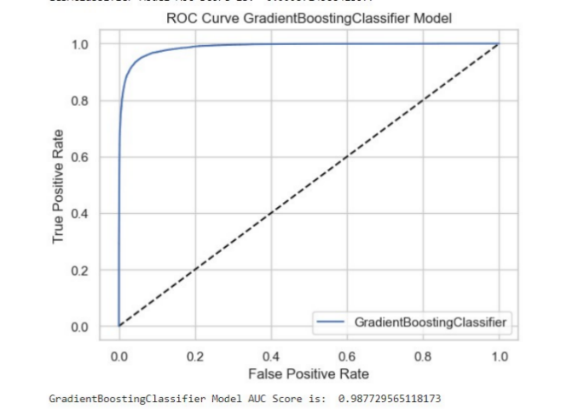
## 

## Figure 23. Light GBM Model Performance First analysis.

The model's stated accuracy is 1.00, indicating exceptional performance. The algorithm discovered 46826 fraudulent transactions out of a total of 47154 while expecting 46898 real transactions out of 47444 expected legal transactions, resulting in 546 misdiagnosed valid transactions. Furthermore, in 328 of the fraudulent transactions, the model was identified as real rather than fake. In general, the model behaved admirably.

"amt" and "log\_amt" are the most relevant features for the model's decision-making component, followed by the rest of the features, indicating that for this model, all features are significant for prediction.

Gradient Boosting Model



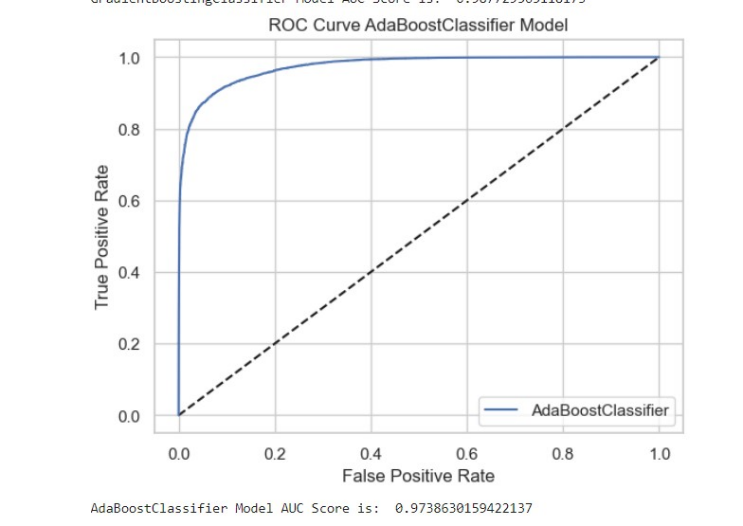
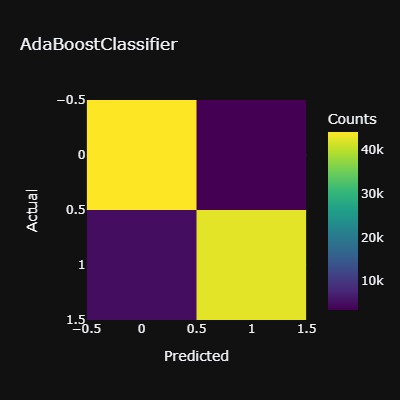
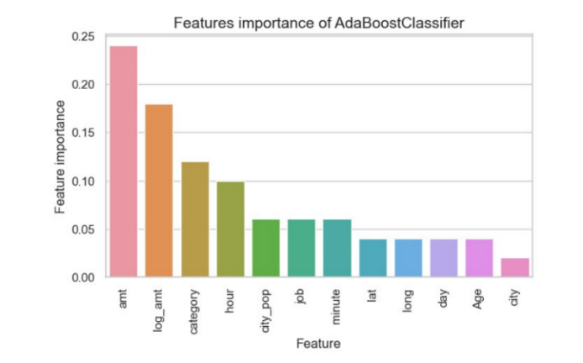
## Figure 24. Gradient Boosting Model Performance First analysis.

The reported accuracy of the model is 0.94, suggesting good performance. The system identified 43765 fraudulent transactions out of a total of 47154 while anticipating 45533 actual transactions out of 47444 predicted legal transactions, resulting in 1911 misidentified valid transactions. Furthermore, the model was identified as genuine rather than fraud in 3389 of the fraudulent transactions. It should be noted that just because a model has a high accuracy does not mean it is reliable; for the goal of credit card fraud, this is not appropriate due to the vast number of fraudulent transactions that could not be identified.

"amt" is the characteristic having the highest relevance for the model's decision-making component, while the rest of the features have extremely little value if any at all.

AdaBoost model

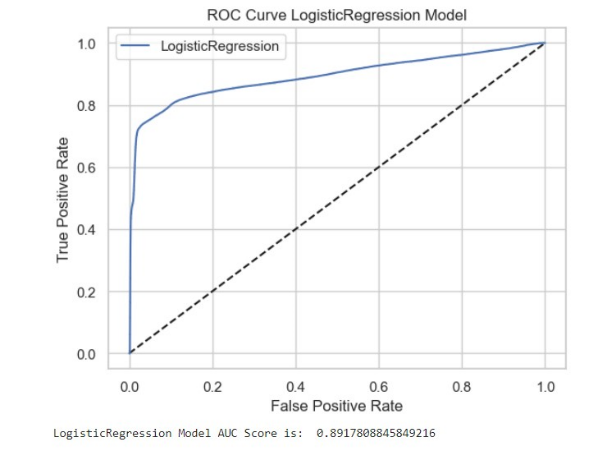
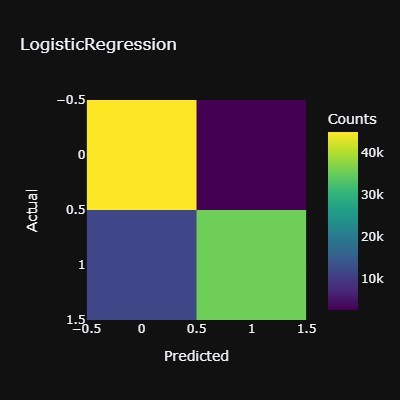
## Figure 25. AdaBoost Model Performance First analysis.



The model's stated accuracy is 0.91, indicating good performance. The algorithm detected 42326 fraudulent transactions out of 47154 total transactions while anticipating 43974 real transactions out of 47444 projected legal transactions, resulting in 3470 misidentified genuine transactions. Furthermore, in 4828 of the fraudulent transactions, the model was determined as real rather than fake. This is not ideal for the purpose of credit card fraud owing to the large number of fraudulent transactions that could not be discovered or were identified incorrectly.

"amt" is the feature with the most significance for the model's decision-making component, followed by log\_amt, and finally all of the other characteristics exhibit importance for model prediction.

Logistic Regression



## Figure 25. Logistic Regression Model Performance First analysis.

The reported accuracy of the model is 0.85, suggesting reasonable performance. The algorithm discovered 35547 fraudulent transactions out of 47154 total transactions while predicting 45030 legitimate transactions out of 47444 projected legal transactions, resulting in 2414 genuine transactions that were misdiagnosed. Furthermore, the model was judged to be actual rather than false in 11607 of the fraudulent transactions. Because of the enormous number of fraudulent transactions that could not be found or were identified wrongly, this is not ideal for credit card fraud.

Results Comparison

A screen shot of a graph

Description automatically generated

## 

## Figure 26. Comparison of the Models Performance First analysis.

Random forest has the highest accuracy at 100%, as well as the highest results in the other parameters, recall, precision F1, and AUC-ROC. It is worth noting that 11 out of 47154 samples were forecasted inaccurately, whereas 116 out of 47744 real transactions were misidentified. In conclusion, the model performed flawlessly and precisely across all metrics and reports examined for modelling and testing; based on this study, it is regarded as the ideal model for detecting credit card fraud in this dataset.

# **4.1.2 Results Second** Analysis

Results after applying Under-sampling for class imbalance.

Following the use of strategies for addressing the class imbalance problem, which is classified into three types: resampling approaches, ensemble-based approaches, and cost-sensitive learning approaches. This dissertation focuses solely on the resampling technique and the ensemble-based approach. Misclassification costs are taken into account in cost-sensitive learning. For financial institutions, for example, the cost of incorrectly projecting a genuine transaction as a fraudulent transaction is significantly higher than the cost of incorrectly projecting a legal transaction as a fraudulent transaction. As a result, by weighting the minority class's misclassification cost more heavily than the majority class's, the model's actual positive rate may be enhanced.

Random Under sampling

In order to balance the dataset, this approach randomly removes the majority samples. This strategy is best used when the training data is really large as is the case in the present dataset. Reduced frequency of majority sampling increases performance while also decreasing storage issues. The problem with utilising such a strategy is that some relevant information may be lost in the process of removing the majority of samples. As a result, the classification prediction may be inaccurate.

Random Oversampling

To balance the dataset, this approach randomly duplicates the minority samples. This method, unlike random undersampling, does not result in information loss. However, because it duplicates the minority samples, there is a considerable risk of overfitting the data. This technique was applied and evaluated in the models but affected the algorithms because of the complexity and the difficulty of handling a large amount of data.

Under-sampling and Oversampling were evaluated for the models, but under-sampling performed better, aimed at this reason the technique was applied for this dataset because it contains millions of rows that after applying oversampling were duplicated. It was also tested by grouping randomly a certain amount of the oversampling samples; however, the results were not the best.

Lasso

This study was also performed feature selection using the lasso technique, which is a tool that helps minimize the cost function. Lasso regression will automatically choose the features that are beneficial to the model, discarding the redundant features. So, the purpose of using Lasso regression for feature selection goals is straightforward: It was applied Lasso regression on the dataset, and only those features that produce a coefficient different from 0 In the output of the feature selections were considered for the model processing. However, the results were satisfactory, and for this reason was decided to keep all the features for the modelling.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Models** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **AUC Score** |
| Random Forest | 0.98 | 0.98 | 0.99 | 0.98 | 0.9982 |
| Gradient Boosting | 0.98 | 0.97 | 0.98 | 0.98 | 0.9973 |
| AdaBoost | 0.95 | 0.96 | 0.94 | 0.95 | 0.9928 |
| Logistic Regression | 0.68 | 0.63 | 0.91 | 0.74 | 0.8374 |
| Support Vector (SVM) | 0.77 | 0.92 | 0.60 | 0.73 | 0.8598 |
| KNeighbors (KNN) | 0.95 | 0.94 | 0.97 | 0.95 | 0.9828 |
| Naive Bayes | 0.60 | 0.86 | 0.27 | 0.41 | 0.7527 |
| Decision Tree | 0.98 | 0.98 | 0.98 | 0.98 | 0.9758 |
| XGBoost | 0.99 | 0.99 | 0.99 | 0.99 | 0.9989 |
| LightGBM | 0.99 | 0.98 | 0.99 | 0.99 | 0.9984 |

*Table 5. Models Comparison – Analysis 2*

Random Forest Model

A graph of a random forest

Description automatically generated

## Figure 26. Random Forest Model Performance Second analysis.

The reported accuracy of the model is 0.98, suggesting reasonable performance. The algorithm discovered 1333 fraudulent transactions out of 1351 total transactions, while predicting 1254 legitimate transactions out of 1278 projected legal transactions, resulting in 24 genuine transactions that were misdiagnosed. Furthermore, the model was judged to be genuine rather than fraudulent in 18 of the fraudulent transactions. Overall, the model performed very well, with just 42 misidentified from the total of 2629 samples.

Gradient Boosting Model

A chart of a graph

Description automatically generated with medium confidence

## Figure 27. Gradient Boosting Model Performance Second analysis.

The model's stated accuracy is 0.95, indicating adequate performance. The algorithm identified 1312 fraudulent transactions out of 1351 total transactions while predicted 1245 valid transactions out of 1278 projected lawful transactions, resulting in 33 actual transactions being misidentified. Furthermore, in 39 of the fraudulent transactions, the model was determined to be real rather than fraudulent. Overall, the model performed well; nevertheless, when compared to random forest, the model fared lower.

AdaBoost Model

A chart of a graph

Description automatically generated with medium confidence

## Figure 28. AdaBoost Model Performance Second analysis.

The reported accuracy of the model is 0.95, suggesting adequate performance. The model recognised 1283 fraudulent transactions out of 1351 total transactions while predicting 1211 legitimate transactions out of 1278 projected lawful transactions, resulting in 67 misdiagnosed real transactions. Furthermore, the model was judged to be legit rather than fraud in 68 of the fraudulent transactions. The model performed well overall; but, when compared to random forest, the model performed with more difficulties.

Logistic Regression Model

A graph of a logistic regression

Description automatically generated

## Figure 29. Logistic Regression Model Performance Second analysis.

The reported accuracy of the model is 0.68, suggesting a poor performance. The model recognised 1211 fraudulent transactions out of 1351 total transactions while predicting only 565 legitimate transactions out of 1278 projected legible transactions, resulting in 713 misdiagnosed real transactions. Furthermore, the model was assessed to be legit rather than fraud in 140 of the fraudulent transactions. The model performed misidentified many of the samples, for this reason, is not a model considered viable for credit card fraud prediction, because of the high accuracy of the prediction that is required.

Support Vector SVM Model

A chart with different colors

Description automatically generated

## Figure 30. SVM Model Performance Second analysis.

The model's stated accuracy is 0.76, indicating an acceptable performance. The model identified 800 fraudulent transactions out of 1351 total transactions while predicting 1196 genuine transactions out of 1278 projected legitimate transactions, resulting in 82 false positives. Furthermore, in 551 of the fraudulent transactions, the model was determined to be legitimate rather than fraudulent. The model misinterpreted many of the samples, particularly a substantial number of fraudulent transactions; as a result, it is not a good model for credit card fraud prediction, due to the high misidentified obtained that financial institutions demand for this purpose especially because represents losses.

KNeighbors KNN Model

A chart of a graph

Description automatically generated with medium confidence

## Figure 31. KNN Model Performance Second analysis.

The model's identified accuracy is 0.95, indicating satisfactory performance. The algorithm identified 1296 illegitimate transactions out of 1351 total transactions while predicting 1201 valid transactions out of 1278 projected legalised transactions, resulting in 77 false positives. Furthermore, in 55 of the fraudulent transactions, the model was determined to be legitimate rather than fraudulent. Overall, the model performed well; nevertheless, when compared to the random forest, the model performed more difficult in terms of credit card fraud prediction.

Naive Bayes Model

A chart with different colors

Description automatically generated

## Figure 32. Naive Bayes Model Performance Second analysis.

The model's identified accuracy is 0.63, indicating inadequate performance. The algorithm identified only 405 illegitimate transactions out of 1351 total transactions while predicting 1243 valid transactions out of 1278 projected legalised transactions, resulting in 35 false positives. Furthermore, losing 946 of the fraudulent transactions, the model was identified to be legitimate rather than fraudulent. Even though the model performed with high precision identifying almost all the legit transactions; nevertheless, when compared to recall which is the result of predicting the fraudulent transactions the performance is very poor, that couldn’t be used for the main purpose.

Decision Tree Model

A chart of a tree

Description automatically generated with medium confidence

## Figure 33. Decision Tree Model Performance Second analysis.

The model's stated accuracy is 0.97, indicating satisfactory performance. The algorithm identified 1310 fraudulent transactions out of 1351 total transactions while predicted 1233 valid transactions out of 1278 projected legitimate transactions, resulting in 45 actual transactions being misidentified. Furthermore, in 41 of the fraudulent transactions, the model was determined to be real rather than fraudulent. Overall, the model performed was acceptable; nevertheless, when compared to random forest, the model performed inferiorly.

XGBoost Model

A chart of a graph

Description automatically generated with medium confidence

## Figure 34. XGBoost Model Performance Second analysis.

The reported accuracy of the model is 0.99, suggesting an excellent performance. The algorithm discovered 1337 fraudulent transactions out of 1351 total transactions, while predicting 1263 legitimate transactions out of 1278 projected legal transactions, resulting in 15 genuine transactions that were misdiagnosed. Furthermore, the model was assessed to be genuine rather than fraudulent in 14 of the fraudulent transactions. Overall, the model performed very well, with just 29 misidentified from the total of 2629 samples. This makes the model suitable for its excellent performance in predicting fraudulent transactions and a precise model in general.

LightGBM

A chart of a graph

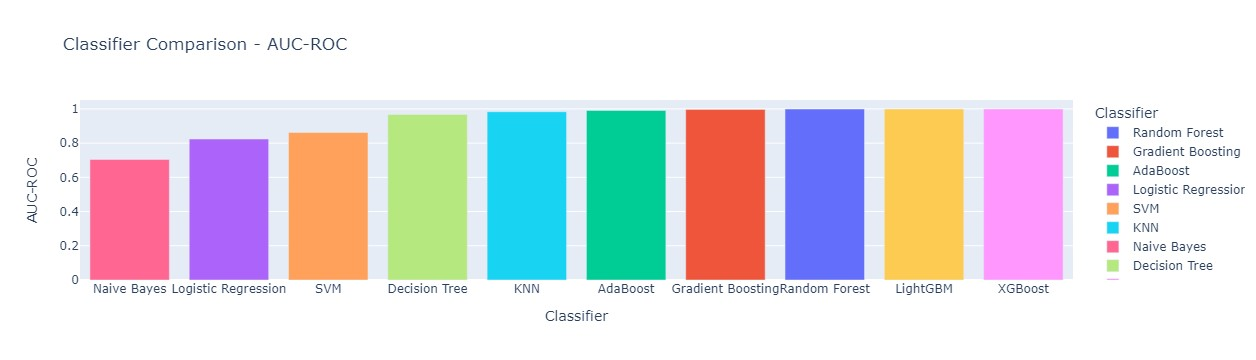
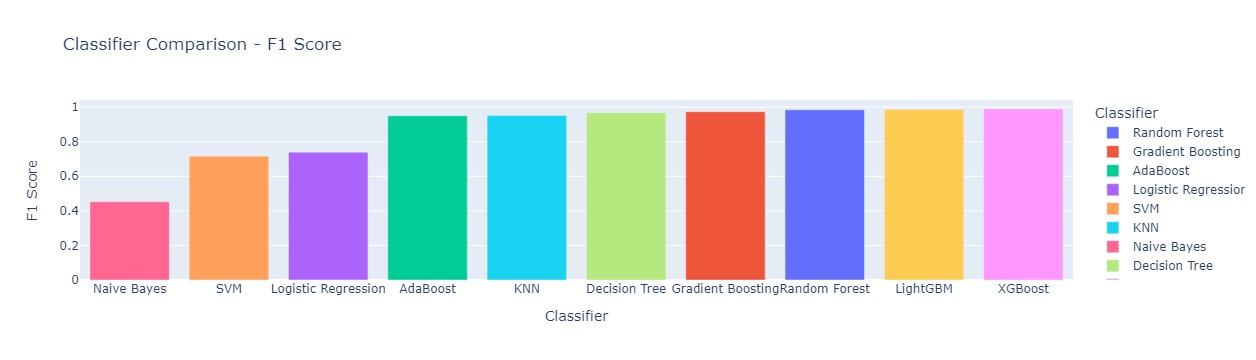
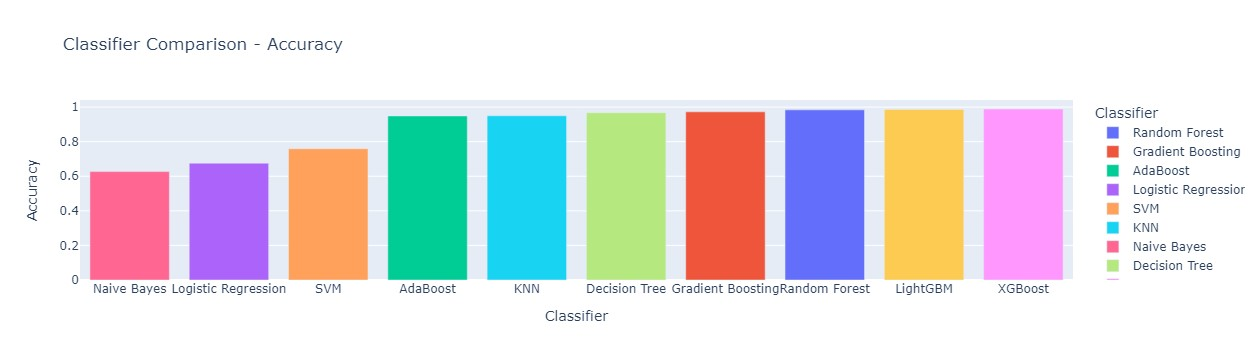
Description automatically generated with medium confidence

## Figure 35. LightGBM Model Performance Second analysis.

The reported accuracy of the model is 0.99, suggesting an excellent performance. The algorithm discovered 1337 fraudulent transactions out of 1351 total transactions, while predicting 1256 legitimate transactions out of 1278 projected legal transactions, resulting in 22 genuine transactions that were misdiagnosed. Furthermore, the model was assessed to be genuine rather than fraudulent in 14 of the fraudulent transactions. Overall, the model performed very well, with just 36 misidentified from the total of 2629 samples. Its performance was exceptional in predicting fraudulent transactions and a precise model in general.

Results Comparison

XGBoost and LightGBM, performed the best in regard to accuracy with 99%, as well as both obtained the same scores for F1 and AUC-ROC, however, there is a slight difference between each other, aimed at legitimate transactions XGBoost, incorrectly predicted only 15 samples, LightGBM inaccurately identified 22 samples, in terms of fraudulent activities both models misidentified the same amount: 14 samples. In conclusion, XGBoost is the suitable model according to the comparison of the models applied for credit card fraud prediction designed for this dataset.



## Figure 36. Comparison of the Models Performance First analysis.

CHAPTER 4

# 4.1 Conclusions and Future Research

In relation to the first dataset, the analysis of various machine learning classifiers has assessed the performance using both cross-validation and test set evaluation. The cross-validation results indicate that most classifiers exhibit consistent and promising performance, with minor variations in mean cross-validation scores. Random Forest and Gradient Boosting demonstrate high mean cross-validation scores of approximately 0.9842 and 0.9777, respectively, showcasing strong generalization capabilities without significant overfitting. LightGBM also stands out with a mean cross-validation score of about 0.9874, indicating robust performance. Other models, including AdaBoost, K-Nearest Neighbors (KNN), and Decision Tree, also exhibit competitive mean cross-validation scores, suggesting their suitability for the task. SVM and Logistic Regression, while having lower mean cross-validation scores, still show consistent results, implying reasonable generalization. Naive Bayes lags with a lower mean cross-validation score of approximately 0.6385 but maintains consistent performance. For the test set evaluation, the classifiers largely maintain their strong performance. Random Forest, Gradient Boosting, XGBoost, and LightGBM continue to excel with high accuracy, F1 scores, and AUC-ROC values, all exceeding 0.98, demonstrating their effectiveness in making accurate predictions on new, unseen data. Despite being a simpler model, Logistic Regression achieves respectable performance with an accuracy of around 0.6755 and an F1 score of approximately 0.7395, along with a reasonable AUC-ROC value of 0.8237. SVM and KNN remain competitive with accuracy values above 0.75 and AUC-ROC values above 0.85, indicating their effective classification ability. Decision Tree exhibits high accuracy, F1 score, and AUC-ROC values, demonstrating its suitability for this task, while Naive Bayes struggles with lower accuracy and F1 scores, suggesting limitations in handling this dataset. In summary, Random Forest, Gradient Boosting, XGBoost, and LightGBM emerge as the top-performing models for this classification task, both in cross-validation and on the test set, offering strong predictive power and generalisation capabilities, making them well-suited for this specific problem.

In consideration of the second dataset, was conducted a comprehensive evaluation of multiple classification algorithms on the analysis. Among the models, Random Forest, Gradient Boosting, and XGBoost demonstrated strong performance, with high mean cross-validation scores and excellent accuracy, F1 scores, and AUC-ROC values on the test set. These models showcase their ability to generalize well to unseen data without overfitting. AdaBoost also exhibited consistent performance, albeit with slightly lower scores compared to the top-performing models. Logistic Regression, SVM, and KNN, while not performing as strongly as tree-based models, still showed consistent and reasonable generalization on both cross-validation and the test set. On the other hand, Naive Bayes, while having lower mean cross-validation and test set scores, maintained consistent performance across different folds. Lastly, Decision Tree performed well with consistent cross-validation scores and high-test set scores, demonstrating its generalization capabilities. The LightGBM model emerged as a top performer, displaying exceptional performance in both cross-validation and test set evaluation, making it a promising choice for this classification task.

After applying all of the methods that were required for each of the analyses carried out for this thesis, it can be identified that even if the goal is the same to achieve every dataset required, it owns experimentation, seeing the goal separately, this is because, in the process of data pre-processing, it can be identified that is highly important each step for the entire analysis, since the moment the data is collected it is important to identify the purpose of the action. This can be concluded because, during the investigation, it was considered the analysis for prediction of fraudulent or legitimate transactions for another dataset that did not contain fraud or legitimate labels but was not possible because it only contained one transaction per day, which according to the analysis is very important for the model to learn and identify patterns related to each transaction.

It is also worth noting that many tactics were taken to generate the greatest metrics report. Although both datasets are severely unbalanced, various techniques were required to cope with the difficulty and get correct findings, as evidenced by the assessment report. All the machine learning models provided valuable insights for credit card fraud prediction, and it is crucial to note that the results produced were sufficient for the main goal, which is to determine the model that best executes credit card fraud detection.

Many of the concepts explored at the start of the thesis have been reduced at the conclusion of the analysis since it strives for a specific analysis to discover what best matches each dataset so that the objective may be reached. This study will assist feature researchers in credit card fraud detection by allowing them to apply their findings and adapt them to the actual world, where the necessity to address the present issue is increasing in today's society.

# 4.2 References

A. Banarescho. “*Detecting and Preventing Fraud with Data Analytics”,* Procedia. Economics and Finance, Vol. 32 (2015) pp. 1827-1836.

A. Patcha, J. M. Park. “*An overview of anomaly detection techniques”:* Existing solutions and latest technological trends, Computer Networks. 2007. pp. 3448-3470.

A. Singh and D. Narayan*. “A Survey on Hidden Markov Model for Credit Card Fraud Detection”,* International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-3, February 2012

Addo, Peter Martey, Dominique Guegan, and Bertrand Hassani*. “Credit Risk Analysis Using Machine and Deep Learning Models”.* 2018, *Risks* 6: 38.

Alaka, Hafiz A., O. Lukumon Oyedele, Hakeem A. Owolabi, Vikas Kumar, Saheed O. Ajayi, Olugbenga O. Akinade, and Muhammad Bilal. “*Systematic review of bankruptcy prediction models”:* Towards a framework for tool selection. Expert Systems with Applications, 2017, 94: 164–84

Alaka, Hafiz A., O. Lukumon Oyedele, Hakeem A. Owolabi, Vikas Kumar, Saheed O. Ajayi, Olugbenga O. Akinade, and Muhammad Bilal*. “Systematic review of bankruptcy prediction models”:* Towards a framework for tool selection. Expert Systems with Applications, 2017, 94: 164–84.

Albanesi, Stefania, and Domonkos Vamossy. “*Predicting Consumer Default*”: A Deep Learning Approach. CEPR Discussion Papers 13914, C.E.P.R. Discussion Papers. London: CEPR. 2019.

Ampountolas, Apostolos, Titus Nyarko Nde, Paresh Date, and Corina Constantinescu. “*A Machine Learning Approach for Micro-Credit Scoring”.* 2021, *Risks* 9: 50.

Bacham, Dinesh, and Janet Zhao. 2017. “*Machine Learning”:* Challenges, Lessons, and Opportunities in Credit Risk Modeling. Moody’s Analytics Report. Moody’s Analytics risk perspectives, Managing disruption. Volume 9, Available online: https://www.moodysanalytics.com/risk-perspectives-magazine/managing-disruption/spotlight/machine-learning-challenges-lessons-and-opportunities-in-credit-risk-modeling (accessed on 3 June 2021).

Barboza, Flavio, Herbert Kimura, and Edward Altman. “*Machine learning models and bankruptcy prediction”.* Expert Systems with Applications: An International Journal, 2017, 83: 405–17.

Bates, John M., and Clive W. J. Granger*. “The combination of forecasts”.* Journal of the Operational Research Society, 1969, 20: 451–68.

Breiman, Leo. *“Arcing the Edge”.* Technical Report 486. Berkeley: Statistics Department, University of California. 1997.

Breiman, Leo. *“Bagging predictors”.* Machine Learning, 1996, 24: 123–40.

Buckland, Steven T., Kenneth P. Burnham, and Nicole H. Augustin. *“Model selection”*: An integral part of inference. Biometrics, 1997, 53: 603–18.

Burnham, Kenneth P., and David R. Anderson*. “Model Selection and Multi-Model Inference”*: A Practical Information-Theoretical Approach, 2nd ed.Berlin and Heidelberg: Springer. 2002.

C. Navamani, M. Phil, 2018. “*Credit Card Nearest Neighbor Based Outlier Detection Techniques”,* International Journal of Computer Techniques. Vol. 5, Issue: 2 (2018). http://www.ijctjournal.org/Volume5/Issue2/IJCT-V5I2P8.pdf.

Chakraborty, Chiranjit, and Andreas Joseph. Working Paper No. 674 *“Machine Learning at Central Banks”.* Bank of England working papers. London: Bank of England, 2017.

Claeskens, Gerda, Jan R. Magnus, Andrey L. Vasnev, and Wendun Wang*. “The forecast combination puzzle”:* A simple theoretical explanation. International Journal of Forecasting. 2016, 32: 754–62.

D. Meyer *“Support Vector Machines”, Technische Universit’’* at Wien, Austria, 2012

Dietze, Michael C. *“Ecological Forecasting”.* Princeton: Princeton University Press. 2017.

E. Duman, and H. M. Ozcelik. *“Detecting credit card fraud by genetic algorithm and scatter search. Elsevier, Expert Systems with Applications”*, 2011; pp13057-13063

E. Ngai, Y. Hu., Y. Wong, Y., Y. Chen, Y. and X. Sun. *“The application of Data mining techniques in Financial fraud detection”*: A classified framework and an academic review of literature,” Decision Support System, pp. 559-569, 2011.

Ewanchuk, Logan, and Christoph Frei. *“Recent Regulation in Credit Risk Management”*: A Statistical Framework. Risks, MDPI, Open Access Journal, 2019, 7: 40.

Fantazzini, Dean, and Silvia Figini. *“Random survival forests models for SME credit risk measurement”.* Methodology and Computing in Applied Probability, 2009, 11: 29–45.

Figini, Silvia, and Paolo Giudici*. “Credit risk assessment with Bayesian model averaging”.* Communications in Statistics. Theory and Methods 2017, 46: 9507–17.

Figini, Silvia, Savona Roberto, and Vezzoli Marika. *“Corporate Default Prediction Model Averaging”*: A Normative Linear Pooling Approach. Intelligent System in Accounting, Finance, and Management. 2016, 23: 6–20.

Granger, Clive W. J., and Ramu Ramanathan. *“Improved methods of combining forecasts”*. Journal of Forecasting 1984. 3: 194–204.

H. Bodepudi, “*Credit Card Fraud Detection Using Unsupervised Machine Learning Algorithms”,* International Journal of Computer Trends and Technology. 2021. https://ijcttjournal.org/archives/ijctt-v69i8p101.

Hand, David J. *“Measuring classifier performance”:* A coherent alternative to the area under the ROC curve. Machine Learning. 2009. 77: 103–23.

Hansen, Bruce E. *“Least squares model averaging”.* Econometrica 2007, 75: 1175–89.

Hastie, Trevor, Tibshirani Robert, and Jerome H. Friedman*. “The Elements of Statistical Learning”*: Data Mining, Inference, and Prediction, 2nd ed. Berlin and Heidelberg: Springer. 2009.

Haykin SS. “Neural networks and learning machines”. 3rd ed. Pearson Education: Upper Saddle River. 2009.

Ian Wright. November 17, 2022. “Credit Card Fraud Capitals of Europe”.

J. Akhilomen. *“Data Mining Application for Cyber Credit-card Fraud Detection System”,* Proceedings of the World Congress on Engineering Vol. III, WCE 2013, July 3 - 5, 2013, London, U.K

J.S Salvatore, W. F David, W. Lee, and L Andreas. Prodromidis; (1997). *“Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results”*; Department of Computer Science- Columbia University, 1997.

K. Chaudhary, J. Yadav, and B. Mallick*. “A review of Fraud Detection Techniques: Credit Card”*, International Journal of Computer Applications (0975 – 8887) Volume 45. 2012, No. 1.

K. Muameleci. “*Anomaly Detection in Credit Card Transactions using Multivariate*”. Generalized Pareto Distribution, Chalmers University of Technology. 2022. https://odr.chalmers.se/handle/20.500.12380/304780

Kruppa, Jochen, Alexandra Schwarz, Gerhard Arminger, and Andreas Ziegler*. “Consumer credit risk”*: Individual probability estimates using machine learning. Expert Systems with Applications 2013 40: 5125–31.

Kuhn, Max, and Kjell Jhonson. *“Applied Predictive Modeling”.* Berlin and Heidelberg: Springer. 2013.

L. Delamaire, H. Abdou, and J. Pointon, *“Credit card fraud and detection techniques”*: a review, Banks, and Bank Systems”, Volume 4, Issue 2, 2009.

L.P. Andreas and J.S Salvatore. *“Agent-Based Distributed Learning Applied to Fraud Detection”;* Department of Computer ScienceColumbia University; 2000.

Lessmann, Stefan, Bart Baesens, Hsin-Vonn Seow, and Lyn C. Thomas. *“Benchmarking state-of-the-art classification algorithms for credit scoring”*: An update of research. European Journal of Operational Research 2015. 247: 124–36.

Li, Guiping. 2022 *“The Prediction Algorithm of Credit Risk of Science and Technology Finance Based on Cloud Computing”*. 1 Jan. 2022: 235 – 251.

Liang, Hua, Guohua Zou, Alan T. K. Wan, and Xinyu Zhang. *“Optimal weight choice for frequentist model average estimators”*. Journal of the American Statistical Association 2011, 106: 1053–66.

M. Ummul Safa, R. M. Ganga, 2019. “*Credit Card Fraud Detection Using Machine Learning”,* International Journal of Research in Engineering, Science and Management. Vol. 2, Issue: 11, (Nov 2019).

Madigan, David, and Adrian E. Raftery. 1994. *“Model selection and accounting for model uncertainty in graphical models using Occam’s window”*. Journal of the American Statistical Association 89: 1535–46.

Malhotra P, Vig L, Shroff G, Agarwal P. *“Long short-term memory networks for anomaly detection in time series”.* 2015

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A survey on bias and fairness in machine learning. 2019.

N. Cristianini and J. Shawe-Taylor *“An Introduction to Support Vector Machines and Other Kernel-based Learning Methods”*. Cambridge University Press, 2000.

Nathalie Japkowicz and Shaju Stephen. “*The class imbalance problem”:* A systematic study. Intelligent Data Analysis, pages 429-449, 2002.

Nelder, John Ashworth, and Robert W. M. Wedderburn. *“Generalized Linear Models”.* Journal of the Royal Statistical Society. 1972, 135: 370–84.

R. Bolton, & D. Hand, *“Statistical Fraud Detection*”: A Review. Statistical Science, 17; 2002 pp 235-249.

R. Patidar, L. Sharma *“Credit Card Fraud Detection Using Neural Network”,* 13-14 May 2011, Jaipur, India, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume1, IssueNCAI, 2011.

S. Ghosh, and D. L. Reilly *“Credit card fraud detection with a neural network”,* Proceedings of the 27th Annual Conference on System Science, Volume 3: Information Systems: DSS, 1994.

S. Maes K. Tuyls, B. Vanschoenwinkel and B. Manderick *“Credit Card Fraud Detection Using Bayesian and Neural Networks"*; Vrije University Brussels –Belgium; 2002.

S. Sorournejad, Z. Zojaji, Reza E. Atani and A. H. Monadjemi *“A Survey of Credit Card Fraud Detection Techniques”*: Data and Technique Oriented Perspective, 2016.

Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. Leroy Miler Jr., and Z. Chen, eds. Climate Change: “*The Physical Science Basis”.* Cambridge: Cambridge University Press. 2007

T. Guo, and L. Gui-Yang *“Neural Data Mining for Credit Card Fraud Detection”*. IEEE, Proceedings of the Seventh International Conference on Machine Learning and Cybernetics; 2008 pp3630- 3634

T. Lane *“Hidden Markov Models for Human/Computer Interface Modeling”.* Proc. Int’ l Joint Conf. Artificial Intelligence, Workshop Learning about Users, 1997 pp. 35-44

T. P. Bhatla*, “Understanding Credits Cards Frauds” Card Business review”* http;//www.tes.com/0\_whitepapers/htdocs/credit\_card\_fraud\_white\_paper\_v\_1.0.pdf, 2013.

Toreini, E. “*The relationship between trust in AI and trustworthy machine learning technologies in Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency”* 272–283.) 2020, et al.

W. Fan, M.Miller, S. Stolfo., W. Lee, and P Chan. *“Using Artificial Anomalies to Detect Unknown and Known Network Intrusions”*, Proc. of ICDM01; 2001 pp123-248.

Yuan, Danny. “*Applications of Machine Learning: Consumer Credit Risk Analysis”*. Cambridge: DSpace, MIT. 2015.

Daniel Nelson 2021, https://www.unite.ai/generative-vs-discriminative-machine-learning-models/

Statista 2023. https://www.statista.com/statistics/1273177/ecommerce-payment-fraud-losses-globally/

Dataset 1 https://www.datacamp.com/workspace/datasets/dataset-python-credit-card-fraud

Dataset 2 https://usmart.io/org/deloitte/discovery/discovery-view-detail/17881c8f-4fa2-4832-99ea-eb2f4033eac2